Towards Ca^{2+} Responsive Reticular Materials

- An Evaluation of Suitable Conditions for Self-Assembly of the Protein S100G

A Master thesis in biophysical chemistry, by

Andreas Carlsson

LUNDS UNIVERSITET

Department of Biochemistry and Structural Biology
Supervisors: Sara Linse and Thom Leiding
Examiner: Mikael Akke

Januari 2021 - June 2021

Abstract

To build materials of proteins is an art that nature is a specialists in. But when we actively try to design protein materials, it becomes clear how complex and hard to control the biophysical world is. This master thesis aims to lay the foundation for a protein material that forms and dissociates as a response to changed Ca^{2+} concentration. If such a material is to be produced with an ordered structure, it might absorb specific molecules, such as a medical substance, and release it under a certain condition. The building block used in this project is a Ca^{2+} transporting protein in our bodies, called S100G. To understand how this protein suitably forms aggregates, the project mainly deals with evaluation of aggregation conditions, to find out when this is happening and for which mutations of S100G it occurs the most. To accomplish this goal, native gel electrophoresis and size exclusion chromatography have been used. The results show that S100G, with an extra inserted linker of nine proline residues between its subdomains (EF-hands), formed three oligomeric structures, in addition to the monomer. The aggregates were stable and only weakly dependent on conditions such as pH, temperature and ionic strength (in non-extreme ranges). However, there needed to be high enough Ca^{2+} concentration present to saturate the proteins. Towards the end of the project, a protein batch from another part of the purification process was used, and these results seem even more promising, indicating that the proteins formed aggregates with radii about 15 times the monomeric radius.

Sammanfattning

Att bygga material av proteiner är en konst som våra kroppar är specialister på. Men när vi medvetet försöker designa proteinmaterial, så blir det tydligt hur komplexa och svårstyrda biofysikens lagar är. Det här examensarbetet har som mål att lägga grunden för ett proteinmaterial som kan skapas eller dissociera som respons på förändrad Ca^{2+}-koncentration. Om ett sådant material kan framställas med ordnad struktur, så är det möjligt att det kan absorbera specifika molekyler, såsom ett läkemedel, och sedan släppa lös detta på en given signal. Byggstenarna som används i det här projektet är ett Ca^{2+}-transporterande protein i våra kroppar, kallat S100G. För att förstå hur detta protein lämpligt kan forma strukturer så ägnar projektet sig till största del åt att undersöka vid vilka förhållanden som aggregering sker, och för vilka mutationer av S100G som detta händer mest. Experiment med native gel elektrofores och storleksseparationskromatografi visar att S100G, med en extra länk med nio prolineresiduer mellan sina subdomäner (EF-händer), bildade tre oligomera strukturer, förutom monomerer. Aggregaten var stabila och endast svagt beroende av förhållanden såsom pH, jonisk styrka och temperatur (i icke-extrema intervaller). Däremot var det nödvändigt med tillräckligt hög Ca^{2+}-koncentration för att mätta proteinerna. I slutet av projektet användes en proteinvariant från en annan del av uppreningsprocessen, och dessa resultat verkar ännu mer lovande då de indikerar att dessa proteiner bildade aggregat med cirka 15 gånger så stor radie som monomererna.

Acknowledgements

I would like to thank my supervisors Sara Linse and Thom Leiding for all their support and company. I really appreciate their dedication and competencies that they shared in the best of ways. Also, to all other in the lab that always kindly listened to my questions and did their best to help, thank you. Last but not least, thanks to Mikael Akke for his role as examiner, and to Ingemar André for his support regarding this matter.

Abbreviations and Notations

- S100G GX or only GX: S100G with X number of glycine residues inserted between the protein´s two subdomains, EF-hand 1 and 2.
- S100G PX or only PX: S100G with X number of proline residues inserted between the two EF-hands.
- A mutation of a protein will be denoted as the one letter code for the wild type residue, followed by the sequence number, followed by the one letter code of the new residue. For instance, if a proline at position 43 is changed to a methionine, this is denoted as P43M.
- MDS: Microfluidics Diffusional Sizing
- SEC: Size Exclusion Chromatography

Contents

1 Introduction

In nature, it is a well known phenomenon that proteins sometimes undergo mutations or other changes leading to non-covalent aggregation into protein clusters. In the example of the linear aggregates called amyloids, this might cause various diseases, such as Alzheimer's and Parkinson's disease (Jaskolski, 2001). But the same principle of spontaneously formed peptide structures could potentially be used to create new materials. With the right design of these materials, they may form under certain circumstances and dissociate under other, in a predicted manner. An example of an application is a material that order itself into a structure that absorbs drug molecules and releases them when the conditions are the right. This will enable controlled and targeted release of the drug, which is highly wanted since the drug potency increases and the side-effects are reduced.

The usage of proteins as building blocks has the advantages over other polymers that it is possible to produce them from highly controlled genes, which gives precise control and knowledge of the molecules. Also, it is possible to mutate the coding gene and get desired changes, for instance to make it absorb another drug. The resulting material can be biocompatible and thus beneficial for *in vivo* applications. Already, it has been shown that homodimers of the tryptophan repressor are able to form an ordered, non-covalently linked, array in alcoholic solution (Lawson et al.,1993). To further make a material suitable for a biologic surrounding, the material needs to be stable in aqueous solutions. Thus, this project will evaluate protein aggregations in aqueous solutions.

One way proteins may aggregate is via *domain swapping*. This means that a part of a protein (domain or subdomain) looses some interactions with the rest of the protein and instead interacts non-covalently with a domain of another protein (Schlunegger et al., 1997). If this occurs in an extended/open ended manner, it may be called runaway domain swapping, leading to the forming of fibrils (Guo and Eisenberg, 2006).

A mutated version of the protein S100G, which is a Ca^{2+} binding and transporting protein in our body, has been showed to undergo domain swapping and form dimers under aqueous conditions (Håkansson et al., 2001). It is therefore believed that also larger aggregates could be generated of variants of this protein, through runaway domain swapping.

The goal of this thesis study is to evaluate under which conditions some variants of S100G aggregate the most or form clusters of discrete sizes. Such knowledge can then be used in future projects to design an aqueous solution, with substrate chains of S100G variants, that will self-assemble into a continuous network, called a reticulum, in response to Ca^{2+} concentration. The principle is illustrated in Figure 1. It is believed that the structure, if designed correctly, could encapsulate other molecules (here called guest molecules). With this, doors to different applications open up, such as absorption, storage and controlled release of the guest molecules. As mentioned above, drugs may be suitable guest molecules. Also catalytic materials to be a potential target, and they can be designed in a similar fashion.

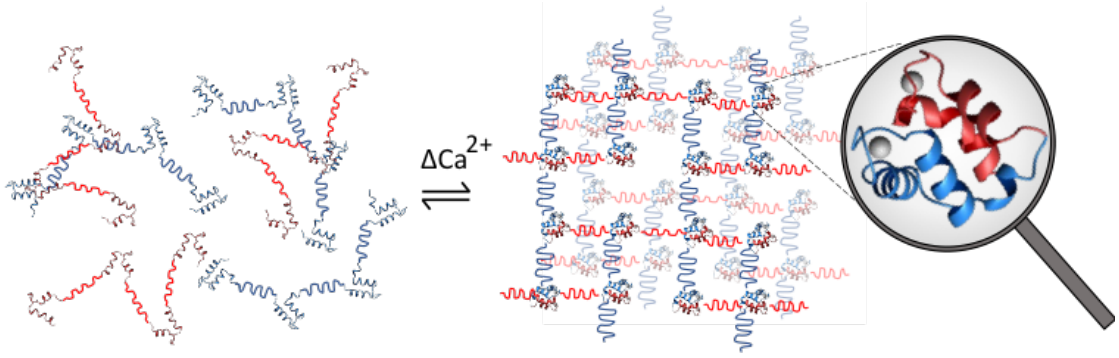

Figure 1: Illustration of how a reticular material might form upon change in Ca^{2+} concentration. The nodes consist of domain swapped, reformed S100G from separate substrate chains. The created voids might get a uniform shape and can then encapsulate guest molecules with affinity for the site. Thanks to supervisor Sara Linse for providing the picture.

1.1 The Protein S100G and Its Protein Family

Human S100G, also named as calbindin D_{9k} and calbindin 3 (with the UniProt ID: P29377) is a Ca^{2+} binding protein in the size of about 9 kDa, or 79 amino acid residues, in the initial form. The protein has two subdomains (motifs) with helix-loop-helix structure (UniProt, 2020; Kördel et al., 1992; Akke et al. 1992)), called EF-hands (Persechini, Moncrief and Kretsinger, 1989). EF-hand 1 (residues 13-48) and EF-hand 2 (residues 45-79) each has a pocket with affinity for a Ca^{2+} (UniProt, 2020). The structure of the protein with and without Ca^{2+} is highly similar (Akke et al. 1991). It is a cytosolic protein, expressed in many different tissues. The role of the protein is not completely understood but it is believed to take up Ca^{2+} in the intestine and transport it in the body, among other (Choi and Jeung, 2008).

S100G belongs to the S100 protein family, which in turn is a part of the super-family of EF-hand proteins. There are about 20 S100 proteins and they are relatively small, in the range of 10-12 kDa. Each contain two EF-hands with affinity for Ca^{2+} and sometimes other divalent metal ions. S100G is monomeric and the other appear as homodimers, heterodimers or higher oligomers (Marenholz et al., 2004). For instance, S100B is shown to form tetrameric, hexameric and octameric structures (Ostendorp et al., 2007; Thulin et al., 2011)), the S100A8/S100A9 heterodimer forms heterotetramers in the presence of Ca^{2+} (Leukert et al., 2006), homodimers of S100A12 are able to arrange (with Ca^{2+}) into spherical hexameric structures in the size of about 55 Å (Moroz et al., 2002) and S100A4 dimers may, in a similar fashion, form a pentameric ring (Malashkevich et al., 2010).

Much of the research about S100G is performed with bovine S100G. Therefore, it is of relevance to know if this knowledge can be used to predict how the human protein will behave (which will be used in this project). A sequence alignment shows an 88.6% identity match between human and bovine S100G. Taking into account similar substitutions, it is as high as 94.9% positive matches. The alignment can be found in Figure 2.

```
P02633 S100G_BOVIN        1    MSAKKSPEELKGIFEKYAAKEGDPNQLSKEELKLLLQTEFPSLLKGPSTLDELFEELDKN    60
                               MS KKSPEELK IFEKYAAKEGDP+QLSK+ELKLL+Q EFPSLLKGP TLD+LF+ELDKN
P29377 S100G_HUMAN        1    MSTKKSPEELKRIFEKYAAKEGDPDQLSKDELKLLIQAEFPSLLKGPNTLDDLFQELDKN    60

P02633 S100G_BOVIN       61    GDGEVSFEEFQVLVKKISQ                                            79
                               GDGEVSFEEFQVLVKKISQ
P29377 S100G_HUMAN       61    GDGEVSFEEFQVLVKKISQ                                            79
```

Figure 2: Sequence alignment of bovine and human S100G, performed in Uniprot with a BLAST search. Identities are shown on the middle line, where + indicates high similarity.

There are four residues that differ significantly, but the first in the sequence, in position 3 (where the bovine type has an alanine and the human type has a threonine) is located in a region that is cut off when the protein is in its Minor A form, and this is the form used in this project. To know where the rest of the substitutions are located in the protein, a PDB file of the bovine variant (PDB entry 2BCB) was viewed, using the protein visualisation program PyMol. According to this, position 45, which has a serine in the bovine protein and asparagine in the human variant, is located in the loop connecting EF-hand 1 with EF-hand 2. But in the protein variant that is used in this project, the mutation N45S is performed and thus, this residue does not differ from the bovine protein. The other two significant substitutions are located at the surface of the protein. At position 12, a glycine in the bovine type is an arganine in the human type. This gives the human type an increased charge and hydrophilicity. Charged side chains at a protein surface may change the stability of the protein (Akke and Forsén, 1990). But, at position 37, the bovine type has a threonine and the human type an alanine, which changes the hydrogen bonding functionality to the opposite. It is hard to predict exactly how the properties change due to this, but since no residues in the interface between the EF-hands differ drastically, the affinity between these can initially be assumed to be closely similar. The only difference in the interior is a leucine residue that is a isoleucine residue in the human variant. These residues are very similar in size and hydrophobicity, but the side-chain structure differs and this might affect the protein structure slightly. Also notable, at position 24, asparagine differs from aspartic acid, and is located close to the Ca^{2+} binding site in EF-hand 1. The more negatively charged site for the human type might increase the affinity towards metal ions. None of the other differences are located close to a binding site.

To summarise, knowledge gained in studies of bovine S100G can likely be translated to the human version and can be used to design initial experimental set ups.

1.1.1 Folding Stability Upon Varying Conditions

In the attempts to understand how a polymerisation of S100G might occur, and what properties can be expected, it is of importance to know how the protein stability and Ca^{2+} affinity are affected when the conditions are changing. Impacts of some parameters have already been studied by various sources and are summarised below.

It shall be noted that in all cited publications in this section, the bovine type of S100G has been studied, but in this project, mutated EF-hands of human S100G is used in addition to some bovine substrates. The reason is that historically much more studies were performed with the bovine variant. Hence, in this section, if nothing else is stated, properties of the bovine S100G are considered.

The wild type of S100G has high affinity between the subdomains EF-hand 1 and 2. Non-covalent interactions, especially hydrophobic interactions, are driving forces. The

dissociation constant, K_D, is about 3 pM in the presence of Ca^{2+}. This is a very strong interaction and the protein does not open up spontaneously on a relevant time scale (half life of years). Even in 10 M urea at 90 C°, the protein does not unfold (Berggård et al., 2001). It can be mentioned that the covalent linker between the EF-hands highly contributes to the total denaturation stability in the absence of Ca^{2+} (Linse et al., 1993).

In a solution with EF-hand fragments of both type 1 and 2, heterodimers are highly preferred to be formed over homodimers (in presence of Ca^{2+}). The heterodimers of the EF-hand fragments are very similar in structure to the intact S100G (Finn et al., 1992; Linse et al., 1993).

Ca^{2+} Concentration
With access to Ca^{2+}, the affinity for the two EF-hands is very high, with a K_D of about 3 pM (Berggård et al., 2001). In absence of Ca^{2+}, this affinity is very much reduced, but still higher than between homodimers of EF-hands fragments (Linse et al., 1993). This tendency is seen in the half life of structure opening, which occurs on the time scale of years in presence of Ca^{2+}, and is reduced to hours in its absence (Linse et al., 1990). For individual fragments of EF-hand 1 and 2, the Ca^{2+} affinity is much lower than the heterodimer/intact S100G. (Finn et al., 1992; Linse, et al., 2020).

The binding of Ca^{2+} occurs with high affinity and positive cooperativity. For the slightly mutated version of S100G, P43M, the cooperativity is measured to 8.5 $kJ{\cdot}mol^{-1}$, corresponding to a 31 times higher Ca^{2+} affinity for binding to the second site when the first site is occupied (Linse et al., 1993).

Temperature
The thermal denaturation midpoint, T_m, for the Ca^{2+}-free (apo) form of S100G, can be estimated to about 85 °C, from experiments using the variant P43M. In complex with Ca^{2+}, the protein is much more stable and is thus hard to measure. T_m of homodimers of EF-hand fragments can be measured in presence of Ca^{2+} to 80 ± 4 °C for EF-hand 1 and 59 ± 2 °C for EF-hand 2 (Linse et al., 1993).

pH
The rate constants of reconstitution of EF-hand 1 and 2 have a very weak pH dependence between pH 5.5 and 9.5. (Dell'Orco et al. 2005). But at pH 5, some important amino acid residues for binding of Ca^{2+} are protonated, leading to reduced Ca^{2+} affinity (Håkansson et al., 2001). Around physiological pH and up to 9, the Ca^{2+} affinity is quite independent of pH and is about its maximum value. But below 7, the dependence is strong and the Ca^{2+} affinity decreases fast with lower pH. Decrease of pH from 7 to 4.5, gives an approximate 10 000 times less Ca^{2+} affinity (Kesvatera et al., 2001).

Mg^{2+} Concentration
Addition of Mg^{2+} induces S100G to be in the Ca^{2+} free form more often. (Håkansson et al., 2001). The macroscopic binding constants of S100G towards Ca^{2+} have been measured by Linse et al., 1991 and Andersson et al, 1997. Corresponding measurements towards Mg^{2+} has been done by Andersson et al, 1997. They are shown in Table 1.

Table 1: Macroscopic Mg^{2+} and Ca^{2+} association constants, K_1 respectively K_2, in mol^{-1}, at low respectively high concentration KCl, from Linse et al., 1991 and Andersson et al, 1997.

KCl concentration (mM)	Mg^{2+} logK_1	Mg^{2+} logK_2	Ca^{2+} logK_1	Ca^{2+} logK_2
<0.2	4.6 ± 0.1	2.7 ± 0.1	8.2 ± 0.1	8.6 ± 0.1
150	3.0 ± 0.1	0.9 ± 0.1	6.3 ± 0.1	6.5 ± 0.1

With considerably higher concentration of Mg^{2+} than Ca^{2+}, as is the case in a resting cell, the site at EF-hand 2 will be occupied by a Mg^{2+} to a large extent. Then, the other site, number 1, is empty since the binding of Mg^{2+} reduces the Ca^{2+} affinity at this site with about a factor of five. But when the concentration of Ca^{2+} increases enough, the empty site is occupied by a Ca^{2+} and this leads to a conformational change, which instead reduces the Mg^{2+} affinity of site 2 about five-fold, leading to an allosteric release of the Mg^{2+}. Now the empty site is rapidly occupied by a Ca^{2+}. Thus, a significant concentration of Mg^{2+} increases the Ca^{2+} cooperativity, but lowers the total Ca^{2+} affinity with approximate a factor of five. Binding of Mg^{2+} seems to change the tertiary structure of the protein less than binding of Ca^{2+} does. (Andersson et al., 1997).

Mutations

Introduction of 1-16 glycine residues in the linker loop between the EF-hands only marginally affects the cooperativity and affinity of Ca^{2+}. But it gives a nearly linear relationship between longer linker and destabilisation of the closed fold, in absence of Ca^{2+}. The average minimum cooperativity for S100G with polyglycine linkers of 1-16 in length was shown to be - 8.6 ± 0.1 kJ/mol and the average binding constants to be 7.79 ± 0.08 for logK_1 and 8.66 ± 0.06 for logK_2 (Linse, et al., 2020).

The mutation P43M is common in many experiments since it makes it possible to cleave the covalent loop between the EF-hands. The resulting protein from the mutation is very similar to the wild type in structure (Finn et al., 1992) and also the thermal stability of P43M is similar to the one for the wild type. (Linse et al., 1993)

Ionic Strength

The two EF-hands, 1 and 2, of S100G have net charges of -1 and -6 respectively. (Finn. et al. 1992). 30 of the amino acid residues are charged, which indicate that the protein will somehow be affected by ionic strength, since this will suppress the charged interactions by screening. In the case of two oppositely charged parts, the screening will reduce the attraction and if they are of the same sign, the repulsion will decrease. Since the Ca^{2+} binds to S100G at negatively charged sites, it is not surprising that the addition of 0.15 M KCl decreases the affinity for Ca^{2+} drastically. The association constants, K (unit M) for site 1 and 2 are logK_1 = 8.2 respectively K_2 = 8.6 with no added KCl, but reduced to logK_1 = 6.3 and logK_2 = 6.5 in the presence of 0.15 M KCl (Linse et al., 1991).

However, the salt dependence for the affinity between the two EF-hands can not be described in this way. They are of opposite signs when binding Ca^{2+} (+1 and -4), but the affinity increases with addition of NaCl, until about 0.15 M (five-fold higher affinity) and then it decreases again with higher concentrations. This might be due to conformational changes at these salt concentrations, leading to optimal interactions between the EF-hands (Dell'Orco et al. 2005).

The hydrophobic effect and hence protein stability is affected by ionic strength and what type of ions that are present. Depending on the size and charge of the ions, they will affect the hydrophobic effect to various extent. The ability to precipitate proteins (salting out/in) follows the Hofmeister series. Also, it depends on the protein properties and the ion-protein interactions (Sun and Sun, 2016).

Denaturants
In the absence of Ca^{2+}, denaturation with urea has been analysed by Linse et al., 2020, showing that denaturation can be described by a two-state process. At slightly higher concentration than 5 M urea, half the population of S100G were denatured. The same assay was performed for mutations with polyglycine of different lengths inserted between the EF-hands. From this, it was noted that the longer the linker, the lower the stability. The protein with the longest linker length (16 inserted glycine residues) had half the population denatured at 4 M urea. (Linse et al., 2020)

Protein Concentration
Acoording to Håkansson et al., 2001, using a concentration of 4.5 mM protein makes it take 30 days for 40% of a slightly mutated S100G to form dimers. An other study, of Linse et al., 2020, used 2.5 mM protein in a test for analysing oligomerisation of S100G with extended linkers of polyglycine. In this case, seven different oligomeric structures were found after 48 days.

1.1.2 EF-hand Swapping

Domain swapping is defined as breaking non-covalent interactions between subdomains within a protein and forming new interactions with a protein close by. Dimers, or oligomers of higher orders, are formed as a result of domain swapping (Schlunegger et al., 1997). For S100G (mutated variants), this is happening via the Ca^{2+} free (apo) form and it is an EF-hand subdomain that finds an other protein´s EF-hand (Håkansson et al., 2001). If the swapping occurs so that each protein finds an adjacent protein and thus extends in an open ended manner, it may be called runaway domain swapping. In this case, longer fibrils are formed (Guo and Eisenberg, 2006).

Since the S100G structure opens up on a timescale of years in presence of Ca^{2+}, swapping will not occur significantly under this condition. Since the corresponding time for the apo form is a few hours, swapping is possible here. The use of sub-saturating concentrations of Ca^{2+} is one way to induce swapping, but there are also other factors that can be altered in combination to this. The strategy is to make the open form more common and this can be done by lowering the pH, addition of Mg^{2+} or via mutations that lower the Ca^{2+} affinity and/or increases the stability of the apo form. Mutations like P43M can slightly stabilize the open state and thus enhance swapping. Also, this mutation will stabilize dimerisation since it enhances a stronger hydrophobic core in the dimeric state (Håkansson et al., 2001.

Once formed, the swapped dimeric structure is more stable since both the proteins need to release Ca^{2+} and open up their EF-hands in order to become monomeric. This results in a significant energy barrier to separate a dimer and rise of temperature to over 45 °C is a way to divide the structures into monomers. A solved structure of bovine S100G, using X-ray diffraction, can be seen in Figure 3. The protein carries the mutations P43M and Q22N. In that study, dimers were found to be formed frequently, but no higher oligomers (Håkansson et al., 2001).

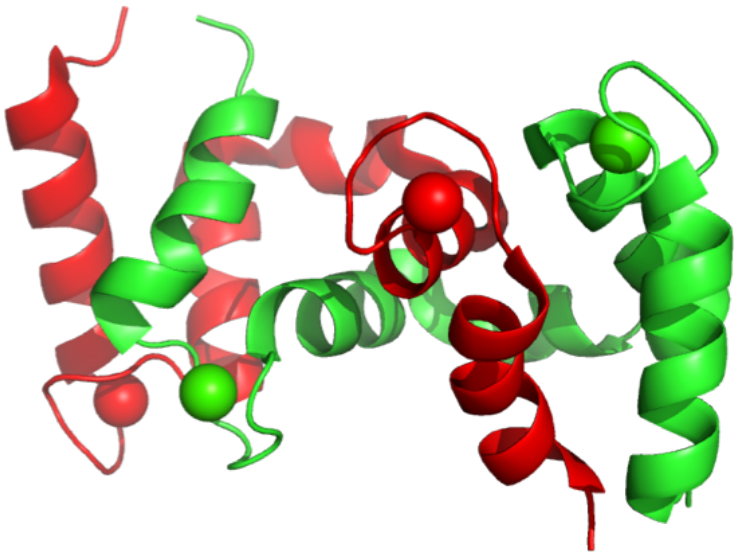

Figure 3: Dimer of bovine S100G, with the two mutations P43M and Q22N. The two protein monomers (coloured red and green) have swapped domains with each other and formed a dimer. The four bound Ca^{2+} are illustrated as spheres. The figure is generated with PyMol, from the PDB-file 1HT9, which is the structure solved by Håkansson et al., 2001, using X-ray Diffraction.

1.2 Topology Analysis

To get an idea of what kind of structures are possible with various substrates, a thought experiment was performed, based on previously findings and with the help of the protein visualisation program PyMol.

With distance measurement tool in PyMol, the following data were obtained:

- A left-handed helix of Gly_8 has a length of 24 Å (PDB file 6f45).
- A left-handed helix of Pro_5 has a length of 14 Å (PDB file 1vzj).
- A left-handed helix of Pro_9 has a length of 24 Å (PDB file 4btb).
- The approximate dimension of the closed protein S100G were found to be 25 Å in diameter (PDB file 2bcb).
- The approximate length of the original linker loop is 25 Å (PDB file 2bcb).
- The approximate dimensions of a S100G dimer is 46 Å x 26 Å x 26 Å (the PDB file 1ht9).

It is worth mentioning that the total linker loop that connects the EF-hands in this study is the inserted linker *plus* the original loop. For example, a substrate with 9 proline residues inserted has a total linker loop of about 25 + 24 = 49 Å. However, whether this loop will be extended mostly or if it will adapt a more rigid fold is difficult to predict. Probably the glycine linkers will be more flexible than the ones of prolines, due to the small side chain of glycine and given the more restricted conformations of proline.

Figure 4 illustrates how monomers, dimers, trimers and tetramers might assemble. It is also possible that runaway domain swapping occur, creating linear chains illustrated in Figure 5. The structures may form annular oligomers, i. e. circles if the two open ends find each other. If the linkers have any preferred angle, the radii of the circles will be discrete in sizes and also helices might be formed. Since each protein only has two sites for other proteins to bind in, and otherwise an inert surface, no predictable three dimensional *network* is likely to be formed. But some mechanically tangled structures might result in solid like behaviour if high enough concentration and long enough chains.

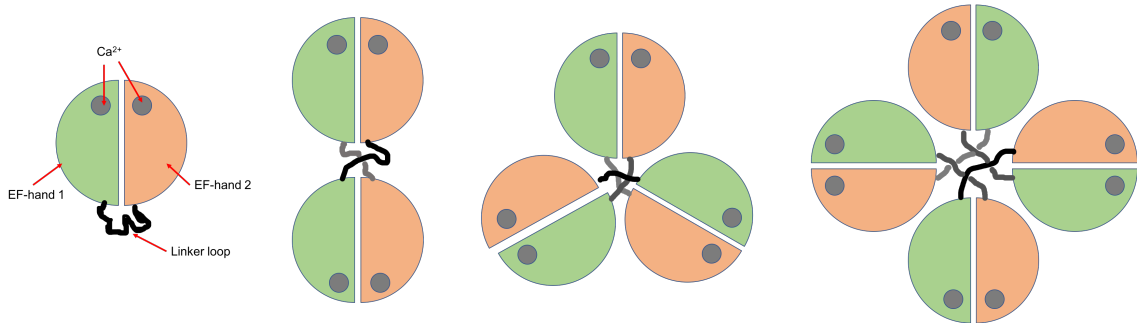

Figure 4: Illustration of S100G monomer (to the left) and the smallest possible aggregates.

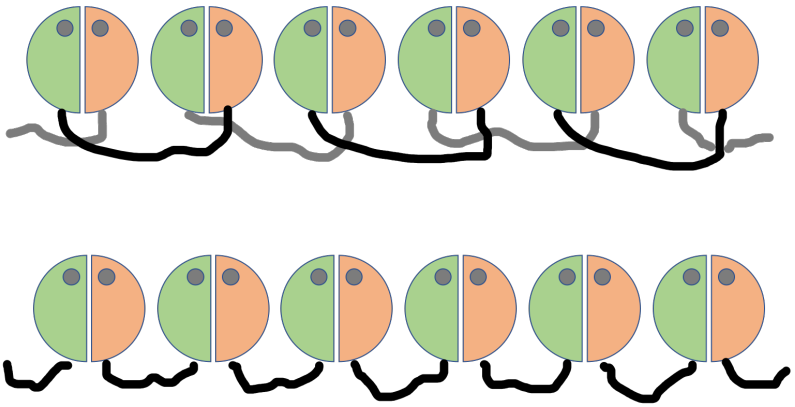

Figure 5: Two examples of possible runaway domain swapping. The first one has intertwined linkers and EF-hands that are not rotated in comparison to the monomeric structure. The second one needs a rotation of the EF-hands of 180°, but the linkers can be shorter or less extended. Depending whether any specific angles are preferred of the linkers, circles or helices of discrete sizes might be likely to be formed.

If each, or some, protein substrates would have three or more points of interactions with other proteins, a network may be formed and give gel-like structure with considerable yield stress (Seiffert, S., 2015). This can, for instance, be gained with substrates of three EF-hands covalently linked, such as EF1-EF1-EF1 in mixture with EF2-EF2-EF2.

Four dimers of S100B may form an octameric structure (Ostendorp et al., 2007). It is possible that domain swapping could be induced in this structure, connecting many octamers together and build up a network.

2 Methods

To accomplish the goal of this thesis, different analytical methods were used. To know what methods are suitable for certain analysis, this project started with a literature search for different possible analytical methods. The ones that became used in the project are described in a general manner at first, and then the specific experimental procedures are given, in the overall order at which the methods were used.

2.1 General Descriptions of Analytical Methods

2.1.1 Absorbance Measurements for Concentration Determinations

When light is traveling trough a medium, not all light will penetrate. The absorbed amount of light depends on the material parameter *extinction coefficient*, and the pathlength the light is traveling in the medium. For a solution with some molecules of interest, it is possible to determine the concentration of these molecules by measuring the absorbance. The absorbance should be measured of a wavelength for which the molecules have a significant extinction coefficient. The relation of Beer-Lambert´s law may be used:

$$A = c\epsilon l \tag{1}$$

where A is the absorbance, ϵ is the extinction coefficient for the molecules to be analysed at the same wavelength which the absorbance is measured, and l is the pathlength the light is traveling through the substance (Swinehart, 1962). All cuvetts used in this project had the pathlength 1 cm.

By subtracting the absorbance of the same measurement, but without the molecules that are to be analysed (that is, absorbance of the cuvette with only buffer solution), it is possible to determine the absorbance of the molecules of interest.

2.1.2 Measuring Ca^{2+} affinity using a chelator

One possible way to analyse the Ca^{2+} affinity of a protein, is to mix the protein with a Ca^{2+} chelator that changes its absorbance when it binds Ca^{2+}. If a solution containing Ca^{2+} is added in small steps and the absorbance is measured after each addition, the affinity of the protein towards Ca^{2+} can be calculated from curve fitting using the equilibrium constant as an adjustable parameter. To gain high precision in the analysis, it is important that the chelator has about the same affinity for Ca^{2+} as the protein has, and that the concentrations are in the same regime. This causes Ca^{2+} to be distributed roughly evenly between chelator and protein, allowing more precise analysis of binding attributes, such as the Ca^{2+} affinity and cooperativity of the protein. Quin-2 is a chelator that may be used. It is a molecule that changes its absorbance spectrum when it binds Ca^{2+}. At 263 nm, the absorbance changes the most (Linse, S., 2002). The extinction coefficient of Quin-2 for 239.5 nm light is 42 000 cm^{-1} M^{-1} and Ca^{2+} dissociation constant is 5.2 nM. (Linse et al., 1987).

To start the measurements at a point where most of the substrate is not bound to Ca^{2+}, a buffer with very low Ca^{2+} concentration (< 1 µM) is needed. To generate this low concentration, it is important to minimise the exposure of potential Ca^{2+}, such as glass. Thus, plastic containers should be used if possible and if this is not possible, for example

when using pH electrode, the part in glass can be soaked for about 15 minutes in 2 mM of the Ca^{2+} binding substance EDTA prior to use (Linse, S., 2002).

In this project, the instrument Probe Drum was used for all absorbance measurements and titrations. This is an instrument that can be programmed to automatically perform both titrations and absorbance measurements (Probe Drum, Probation Labs Sweden AB).

2.1.3 Dynamic Light Scattering

Dynamic light scattering, DLS, is used for analysing dynamics of macromolecules in a solution. The diffusion coefficient of the particles can be obtained, and from this, the hydrodynamic radius can be calculated by assuming some shape, e.g. spherical. In the method, monochromatic light is illuminating a solution with particles and intensity fluctuations in scattered light is detected. The fluctuations arise since light will be scattered by the objects in the solution and give rise to destructive respectively constructive interference. When a particle moves due to Brownian motion, the scattered intensities change. From this, it is possible to determine how fast the particles are moving in the medium and gain their translational diffusion coefficient, D_τ. Using Stokes-Einstein equation and with knowledge of the temperature, T, and viscosity, η, of the solution, the hydrodynamic radius, R_H, can be determined for assumed spherical objects (Stetefeld et al., 2016).

$$D_\tau = \frac{k_B T}{6\pi\eta R_H} \tag{2}$$

The raw data of DLS measurements are reported as a so called autocorrelation function, plotted against a delay time. The amplitude of the autocorrelation function depends on the intensity fluctuations (which are specific for each instrument and not very relevant of the size determination) and the delay time is the time between the intensity measurements. The autocorrelation function will decay towards a value of 1, since the particles diffuse and thus change the scattering pattern more and more compared to the initial measurement. The decay is exponential and relates to the size of the particles. (Wyatt Technology, 2014).

DLS is a fast and relatively inexpensive method of studying particles in solution. But the signal is very sensitive to size (Bootz et al., 2004), since the light scattering changes with particle radius to the power of six (Melnichenko, 2016). Hence, if the sample has some dust contamination or other larger particles than the analyte, the signal will be strongly affected and not as reliable. (Bootz et al., 2004). It is a useful method to detect presence of larger particles, since these will be over-represented in the recorded data. (Stetefeld et al., 2016). For monodisperse samples, DLS is accurate and can detect particle sizes of about 1 nm to 1000 nm (Filipe et al., 2010). For the instrument used in this project, a DynaPro PlateReader, the limit of detection is a protein concentration of 0.125 mg/ml for a 14 kDa protein, Lysozyme (Wyatt Technology).

2.1.4 Native Gel Electrophoresis

In an electrophoresis measurement, proteins are placed in a gel with an applied voltage. With a gel of polyacrylamide, the technique is called polyacrylamide gel electrophoresis, or PAGE for short. The proteins are accelerated by the electric field and will move in the gel according to their charge. The gel will make larger proteins move more slowly and hence, the migration speed is determined by the charge and hydrodynamic radius of the protein. Different variants of PAGE are possible and might cause separation based on

different properties. A common approach is to denature the proteins with the negatively charged detergent sodium dodecyl sulfate (SDS), mainly to attenuate the shape and to some extent inherent charge as separation parameter (ThermoFisher). In a native PAGE, the proteins are not denatured and may keep their fold and non-covalent associations. Thus, in a native PAGE, supra-molecules such as dimers or higher non-covalent oligomers might remain intact, which is not the case in a SDS-PAGE (Li and Arakawa, 2018).

2.1.5 Microfluidics Diffusional Sizing (MDS)

In the case of two laminar parallel flows, the fluids will move along the capillary without any mixing beside from diffusion. The fundamental principle of MDS is to measure how much diffusional mixing that occurs of certain particles in such a system. Given knowledge of the capillary dimensions and the flowrate, the translational diffusion constant can be calculated, which relates to the hydrodynamic radius of the diffused particles according to Equation 2, Stokes-Einstein. In this project, the technique will be used with an instrument called Fluidity One, by Fluidic Analytics. Here, 5 µL of a fluid can be added to a chip and the instrument will pump this fluid through a capillary with a parallel laminar flow of an auxiliary stream. In the end of the capillary, the two laminar flows are separated into chambers and a fluorescent reagent is added that react with primary amines, which are present at proteins. The fluorescence is measured and relates to how much diffusion that have occurred over the capillary. The operating protein concentration range is 10-500 µg/mL of the protein Bovine Serum Albumin, BSA, and the size range is 0.5-20 nm in radius. (Fluidic Analytics).

2.1.6 Size Exclusion Chromatography (SEC)

The concept of chromatography is that particles in a fluid (the mobile phase) flows through a stationary phase of some sort, often located in a column. This causes some particles to move slower than other depending on the specific interactions between mobile and stationary phase, and hence a separation is provided. The eluate is the liquid that passes the column and it can be analysed with e. g. absorbance measurements and electrophoresis of collected fractions, to obtain information of the composition. More specifically, size exclusion chromatography uses a stationary phase of gel-beads that causes smaller particles to take a slow way through them and larger particles will move a more straight, faster way through the column. Particles too large to be separated with a certain column will be eluted in the so called void volume.

Fast Protein Liquid Chromatography, FPLC, is a variant specifically developed for separating proteins fast and with high resolution. The stationary phase can be changed to provide separation possible with the most common chromatography techniques, such as affinity, ion exchange and size exclusion. The method uses lower pressure than High-performance liquid chromatography (HPLC) and can also be used for other substances than proteins (Madadlou et al., 2016).

2.2 Experimental procedure

2.2.1 Determination of Substrate Concentrations and Ca^{2+} Affinity Constants

Extraction of a Ca^{2+}-Free Buffer
The procedure followed the description given by Linse, S., 2002, but since this is an important part of the project, a detailed methodology description is given here.

The strongly Ca^{2+} binding Chelex-Resin 100 was dissolved in a plastic container with about 5 liter Millipore water and the pH was adjusted to 7.9. The solution was then vacuum filtered and rinsed several times with Millipore water to reduce any salt. The resin-mass was filled into a dialysis slang with pore-size 6000-8000 kDa, which was firstly cleaned 3 times in boiling Millipore water. A 2 mM Tris buffer was made in a plastic flask and the pH set to 7.5. The Chelex tube was put in the buffer and left to soak Ca^{2+} for three days before measuring the Ca^{2+} concentration. After this, the Chelex tube was kept in the buffer.

To determine the Ca^{2+} concentration of the buffer, the Ca^{2+} chelator Quin-2 was used. Freeze-dried Quin-2 was dissolved in a small volume of the Ca^{2+} free buffer and the Quin-2 concentration was determined with absorbance measurements and Beer-Lambert´s law. $CaCl_2$ was added to gain a concentration of 2 mM, making the proteins well saturated with Ca^{2+}. The blank measurement only contained the Ca^{2+}-free buffer, to secure that only Quin-2 contributed to the absorbance.

The Quin-2 concentration, $[Q]$, was determined to 28.2 µM with the absorbance measurements. With this, the Ca^{2+} concentration of the buffer could be obtained. Absorbance at 263 nm was measured for three situations. Firstly with 28 µM Quin-2 in 1 ml of the buffer, giving the absorbance $A_1 = 1.088$, then, after adding 2.5 µL 0.1 M EDTA to remove any present Ca^{2+} from Quin-2, giving the absorbance $A_2 = 1.097$. Lastly, after also adding 2.5 µL 1.0 M $CaCl_2$, to saturate both EDTA and Quin-2, the absorbance of $A_3 = 0.22$ were measured. See Appendix A, Figure 39, for the absorbance spectra. The initial Ca^{2+} concentration, $[Ca^2+]$ may now be calculated to 0.3 µM with the following formula, from Linse, 2002:

$$[Ca^2+] = [Q]\frac{A_2 - A_1}{A_2 - A_3} \tag{3}$$

Since the Ca^{2+} concentration of the buffer was well below 1 µM, it may be considered suitable for analysing Ca^{2+} binding to S100G and it will from now on be referred to as a Ca^{2+}-free buffer.

Ca^{2+} Titrations of the Substrates
Ca^{2+} titrations were performed as described in detail in Linse. S., 2002, with the difference that smaller cuvetts were used, meaning that volumes were scaled accordingly in the titrations. The titrand volume was 1 ml and volumes of 1.5 µL 3 mM Ca^{2+} solution were added in 30 steps, during constant stirring. The cuvette was always cleaned with EDTA, Millipore water and lastly ethanol, and then dried with nitrogen gas. The blank for all titrations was the Ca^{2+} free buffer, to make sure that only the protein and/or Quin-2 contributed to the recorded absorbance spectra. For protein substrates with polyglycine linkers, the equilibrium time (the time between addition of Ca^{2+} solution and absorbance measurement) was 30 seconds, and for the ones with polyproline linkers, it was 45 seconds. The change was due to an observation that the absorbance might not have the sharp kink

at the point of saturation that was desired.

Firstly, a Ca^{2+} titration was performed with only Quin-2 in Ca^{2+}-free buffer, without any protein. This was done to analyse the linearly decreasing absorbance at 263 nm, when more Ca^{2+} are bound, until saturation.

For determination of the Ca^{2+} affinity constants of the proteins, Ca^{2+} titrations to about 28 µM of both protein and Quin-2, in Ca^{2+}-free buffer, were made. The absorbance data of 263 nm was exported to the software CaLigator (André and Linse, 2002), to do the parameter fittings and gain both concentrations of the proteins and their Ca^{2+} affinity constants.

From the binding constants, the cooperativity can be estimated to a lower limit using the formula $-\Delta\Delta G_{min} = RTln(4K_2/K_1)$ (Linse et al. 1991).

2.2.2 Dynamic Light Scattering for the Evaluation of Aggregation

DLS was used to analyse if aggregation of S100G P5 occurred, at different levels of Ca^{2+} saturation. Freeze-dried P5 was dissolved in 2 mM TRIS buffer, pH 7.5, in four different Ca^{2+} saturation series: 120%, 90%, 50% and one in the Ca^{2+}-free buffer (close to 0% saturation). In each Ca^{2+} concentration series, also four different protein concentrations (2.5 µM, 25 µM, 250 µM, and 2.5 mM) were made, using serial dilution with the Ca^{2+}-free buffer. The samples were made in singlicates.

A volume of 80 µL of each sample was added to a 96-well polystyrene plate (Corning, nr 3880). About 2 hours after dissolutions, the samples were analysed using DLS. The temperature was 25 °C in the machine, the wavelength of the laser was 821.5 nm, acquisition time was set to 5 seconds and 10 acquisitions were made for each well. After the run, it was discovered that each well had a small bubble in the edge of each well, due to bad wetting of the liquid to the polystyrene surface. Thus, the next day, the samples were transferred to a PEGylated plate (Corning, nr 3881). No bubbles were visible and a new run with the same settings were made, this was 24 hours after dissolutions. The plate was always kept at room temperature between the measurements.

In addition to the samples, one well only contained the Ca^{2+}-free buffer, as negative control. This gave strong indications that the buffer contained larger, unwanted particles. Hence, filtration of all samples were made the following day (48 hours after the dissolutions) before a new DLS analyse was done. The filters used had a pore size corresponding to 100 kDa globular proteins.

After five respectively six days had passed since the dissolutions, the plate was analysed in the DLS instrument. No external conditions changed during this time and the samples were untouched.

2.2.3 Native Gel Electrophoresis for the Evaluation of Aggregation

The same samples as used in DLS analysis were used to screen protein concentrations and Ca^{2+} saturation. In other words, P5 were used at Ca^{2+} saturation levels of 120%, 90%, 50% respectively in Ca^{2+}-free buffer, and with protein concentrations of 2.5 mM, 250 µM, 25 µM and 2.5 µM. The samples were mixed with a coloured (coomassie) loading buffer at ratio 1:1, to make the execution of the electrophoresis easier/more visible. 10 µL of each sample with loading buffer was inserted in each well of a 4-20% Tris-glycine gradient gel (Invitrogen, Novex, 1 mm x 15 wells) and a voltage of 150 V was applied. When the blue

coomassie reached the end of the gel, the electrophoresis was ended and the gels put in Instant Blue stain over night, to make the proteins visible.

The same procedure was applied in the coming experiments, where a screening of substrates over time was performed. This time, the protein concentration was the same in all samples; 1 mM. Two levels of Ca^{2+} saturation were used; 80% and 120%, at pH 7.5 and 0.02% NaN_3 as preservative. Fractions of the samples were used after one, two, four, seven respectively fourteen days of incubation at room temperature. The first gel used was of 4-20% gradient Tris-glycine gel and Instant Blue was used to staining over night. But for the following analysis, this was changed to 14% Tris-glycine gels and the staining method Coomassie Brilliant Blue R 250, for about 20-30 min and then destained in 30% ethanol and 7% acetic acid until desired colour of the gel was reached. The reason for the changes was that the 14% Tris-glycine gel was believed to give better separation in the size interval of the aggregates. Similarly, the new staining method was believed to provide more sensitive staining.

In the screening, the substrates P5, P7, P9, G5, G7, G12 were analysed separately and also mixed in various combinations. The combinations were chosen with the intention to try as many logical combinations as possible that could fit in two gels. Thus it was decided to combine the smallest linker in the G-series together with the longest in the P-series, and vice versa (G5+P9 and G12+P5), as well as the substrates with the most similar lengths (G12+P9, G7+P7, G5+P5). Within the same series of linker types, the longest and the shortest were mixed (G5+G12 and P5+P9), as well as all three lengths (G5+G7+G12 and P5+P7+P9). All combinations were dissolved at 120% Ca^{2+} saturation, and all but the last two combinations were also dissolved at 80% Ca^{2+} saturation. The almost unmodified version of S100G, P43M, was present at each gel, to have a reference level of monomers. The S100G was dissolved to 1 mM protein concentration in 6 mM $CaCl_2$, except for the first gel, where the mistake was made to only have the protein concentration 0.1 mM (this was hardly visible so it was changed to the next gel).

The substrate P9 was used for many of the coming experiments in this project and eventually it ran out. Thus, more P9 was desalted and freeze-dried by supervisor Sara Linse. The proteins were from the same expression batch but in another peak of the ion exchange purification. To find out if this new P9 behaved in the same manner as the previously used P9, one more native PAGE was set up, were both the "new" and "old" P9 was used, pH 7.5 and Ca^{2+} saturation 120%.

In the same gel, a series of some more Ca^{2+} saturation levels were evaluated to see if a saturation level higher than 120% would be beneficial for aggregation. Hence, one aliquot of the old P9 was dissolved into three different Ca^{2+} saturation levels: 120%, 150% and 200%. The conditions were otherwise identical, with pH 7.5 and incubation at room temperature for one hour before insertion in the gel wells.

The rest of the wells in the gel were filled with some of the samples used in the earlier performed native PAGE assay, in which the substrates and combinations were screened. These samples had been resting at room temperature for 50 days since the dissolutions (Ca^{2+} saturation in brackets): G5 (80%), G7 (80%), G12 (80%), G5+G7+G12 (120%), G12+P9 (120%), P5 (120%), P5+P7+P9 (120%), P5+P9 (120%), G12+P9 (80%) and P5 (80%).

The gel was a 14% Tris-glycine gel and it was also run at 150V for about 75 minutes and then stained with Coomassie Brilliant Blue R 250.

2.2.4 MDS for Size Determination of Aggregates

Freeze-dried P5 and P9 were each dissolved to a concentration of 1 mM in 2 mM HEPES-buffer, pH 7.5. Three levels of Ca^{2+} saturation levels were used; 80%, 120% and in Ca^{2+}-free buffer. After 24 hours at room temperature, each protein solution was diluted 1:25, to gain a final concentration of about 40 µM, or about 400 µg/ml. 5 µL of each protein solution were added to the inlet of the chip, before starting the run. Also, S100G (P43M) without any inserted linker, at similar concentration and in the HEPES-buffer, was measured in the same way. Each solution was run in triplicates.

One week after the freeze-dried proteins had been dissolved to the various Ca^{2+} saturation levels (stored in 1.5 ml plastic tubes at room temperature), the same procedure of measurements were performed. The sole exception was that P9 in the Ca^{2+}-free buffer was not measured since it had evaporated.

In an other assay, but under the same conditions, also P7 dissolved in Ca^{2+}-free buffer was analysed after 1 day of incubation. MDS was also used in combination with SEC, as described below.

2.2.5 SEC for the Evaluation of Aggregation

A Superdex 75 Increase 10/300 GL column, was used to separate the aggregates based on size. The flow-rate was 1 mL/min and a fraction collector was used to automatically collect 0.5 mL fractions of the eluate in each run. Also in common of all the experiments performed with SEC, is that the loop was washed with at least three times the loop volume with running buffer prior to insertion of sample.

Evaluation of Time Dependence and Stability of Aggregation

Freeze-dried proteins of P5, P7 and P9 were dissolved in 2 mM HEPES-buffer to 1 mM protein, ph 5.7 and at 2.4 mM $CaCl_2$ (120% saturation). Immediately after, 50 µL were inserted in the loop of the FPLC instrument and a SEC program was started that injected the protein solution together with the running buffer. Absorbances at the wavelengths 214 nm, 260 nm, 280 nm and 405 nm were measured of the eluate. The running buffer had the same composition as the one the proteins had been dissolved in, except for 0.1 M NaCl that were added to the running buffer in order for the column to work as predicted. The measurements where repeated in the same way for P5, P7 and P9 at different times after dissolutions. The times were: within a few minutes, after one day, two days respectively seven days. Also a run was performed with newly dissolved S100G (mutation P43M) to have as a reference for a known monomeric structure of about the same size. It was dissolved in the same buffer as the substrate.

To get an indication of the association/dissociation kinetics of the aggregates, the collected fractions of the run with P5 were re-analysed using the same column. 100 µL from the fraction containing the first peak was analysed about 30 minutes after the first run. After an additional 30 minutes, the fraction containing the second peak (of totally three significant ones) was treated in the same way.

In the process of learning how to design and perform the SEC assays, the first run did not have any added Ca^{2+} in the running buffer. This experiment turned out to be of value later. It was performed in the same manner as described above, right after dissolution of freeze-dried protein of P5, P7 and P9. The only difference is that the running buffer did not contain any added Ca^{2+}. In this assay, reruns of fractions were done for P9, with

fractions of both peaks of interest. About 30 minutes after the first run, 100 µL of the first peak (largest particles) were injected, and after an additional 30 minutes also 100 µL of the second peak were injected.

Evaluation of Temperature Dependence of Aggregation

To see if any other equilibrium of aggregation could be reached by varying the temperature, an assay with incubation of 1 mM P9 in 2 mM HEPES buffer, pH 7.5, at 120% Ca^{2+} saturation, was performed. The protein samples were incubated at 4°C, 37°C, 50°C and 70°C directly after dissolutions of the freeze-dried protein. After one hour respectively after one day, SEC was performed with 50 µL of each sample. The running buffer was 2 mM HEPES, pH 7.5, with same concentration of Ca^{2+} and 0.1 M NaCl. The absorbance of the eluate was measured for 214 nm, 260 nm, 280 nm and 405 nm wavelengths.

Evaluation of pH and Ionic Strength Dependence of Aggregation

Freeze-dried proteins of P9 were dissolved to 1 mM at 120% Ca^{2+} concentration and at four different pH, 5, 6 7.5 and 9. For pH 5 and 6, the buffer was made with 2 mM MES monohydrate, for pH 7.5 it was 2 mM HEPES and for pH 9 it was 2mM Tricine. For each pH, two samples were made, one with and one without 0.15 M NaCl. After one hour respectively one day in 50°C incubation, 50 µL of each sample were analysed using SEC, and absorbance of 214 nm and 280 nm was measured of the eluate. The running buffer was the same as in previously assays: 2 mM HEPES, 2.4 mM $CaCl_2$, 0.1 M NaCl and pH 7.5.

To evaluate if the pH-assay provided results that could be used to draw conclusions regarding pH dependence of aggregation, one sample of P9, pH 5 and incubated at 50°C for 1h, was run in a running buffer with pH 5 (with 2 mM buffer MES monohydrate, 2.4 mM $CaCl_2$ and 0.1 M NaCl).

Comparison of Newly Desalted P9 with Earlier Used P9 Aliquots

The protein P9 eluted in two peaks from ion exchange chromatography in presence of Ca^{2+}, the first peak was used initially, and the second peak towards the end of the project. This latter batch of P9 was desalted and freeze-dried by supervisor Sara Linse. Then, the proteins were dissolved in 2 mM HEPES buffer, pH 7.5, at 120% $CaCl_2$ saturation and incubated at room temperature for one day. Under these conditions, a SEC analysis was performed with first a Superdex 75 Increase column, then a Superdex 200 Increase and lastly a Superose 6 Increase. The columns were changed since some particles were too large to be separated. With the Superdex 200 Increase column, 100 µL of a fraction corresponding to the largest particles, repectively 500 µL of a fraction of the belly between this peak and the other peaks, were re-analyzed about 3 hours later. For both re-runs, a 500 µL loop was used.

The particles eluted in the void volume, from the Superose 6 Increase column, were analysed with an absorbance measurement of a broad range of wavelengths. The same measurement was done to the eluate corresponding to short oligomers. In each case, the running buffer used in the SEC analysis was used as blank measurement.

3 Results and Analysis

Each assay section will be introduced by one or a few questions that the assay in question had as main aim(s) to answer.

To generate all plots presented in this report, the software MATLAB has been used.

3.1 Determination of Substrate Concentrations and Ca^{2+} Affinity Constants

Does the Ca^{2+} affinity of S100G significantly change depending on if the inserted linker consist of prolines or glycines? This knowledge was important for designing the further assays and to interpret the results with higher certainty.

The absorbance spectra of Ca^{2+} saturated Quin-2 (see Figure 6) was used to calculate the Quin-2 concentration. The absorbance at 239.5 nm (data point for 239.6 nm) is 1.185 and this gives the concentration 28.2 µM by using Beer-Lambert´s law and 42 000 cm^{-1} M^{-1} as extinction coefficient.

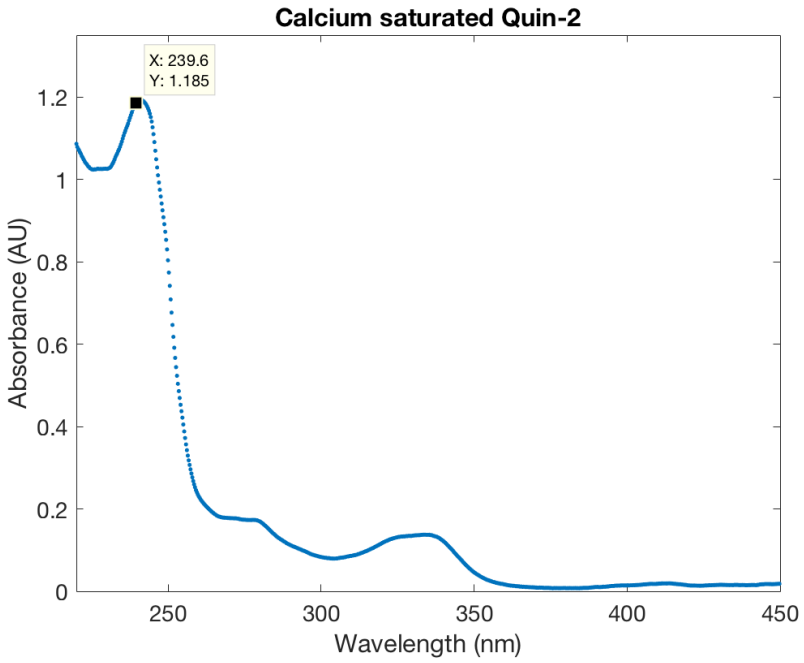

Figure 6: Absorbance spectrum of Ca^{2+} saturated Quin-2 in 2 mM TRIS buffer, pH 7.5. The absorbance for the wavelength closest to 239.5 nm is marked in the graph, 1.185. The sample was measured with the Probe Drum instrument with buffer as blank measurement.

From the Ca^{2+} titration of Quin-2 in Ca^{2+}-free buffer and no protein, the absorbance at 263 nm is plotted using the software CaLigator (André and Linse, 2002) and can be seen in Figure 7. By fixing the protein concentration to 0, the chelator K_D to 5.2 nM and a Ca^{2+} concentration of 3.0 mM in each addition, a fitting was made (see Figure 7). This gives a Quin-2 concentration of 33.6 µM.

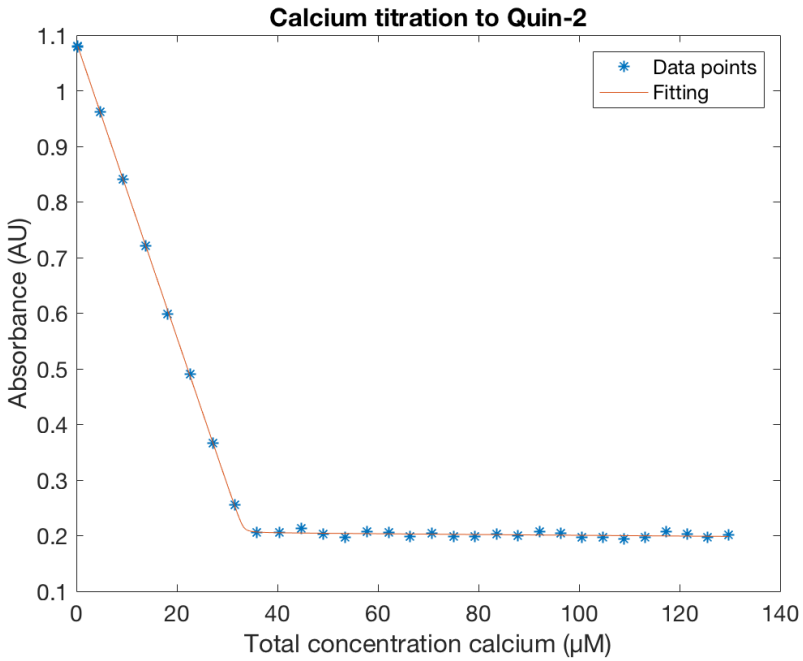

Figure 7: The absorbance of the wavelength 263 nm for about 30 µM Quin-2 in 2 mM TRIS buffer, pH 7.5, is plotted to the total concentration Ca^{2+} during the titration. The parameter fitting from CaLigator is also showed.

Since it is the molar relationship between Ca^{2+} and Quin-2 that determines the absorbance, one can either calculate the Quin-2 concentration from a given amount of added Ca^{2+}, as above, or calculate the Ca^{2+} concentration of the additions from a given Quin-2 concentration. If the latter approach is used, the Ca^{2+} concentration can be calculated using the relation between the Quin-2 concentration given from the absorbance measurement and the Quin-2 concentration corresponding to 3.0 mM Ca^{2+}: $28.2/33.6 \cdot 3.0$ mM ≈ 2.5 mM.

Since the concentrations of Quin-2 and Ca^{2+} affect what proteins concentrations are generated in the fittings by CaLigator, fittings with both concentration pairs were made for each tested substrate. But since the shapes of the data points are the same in these two setups, only the graphs with the higher concentrations are showed here, see Figure 8-9. The other can be found in Appendix A, Figure 40-41.

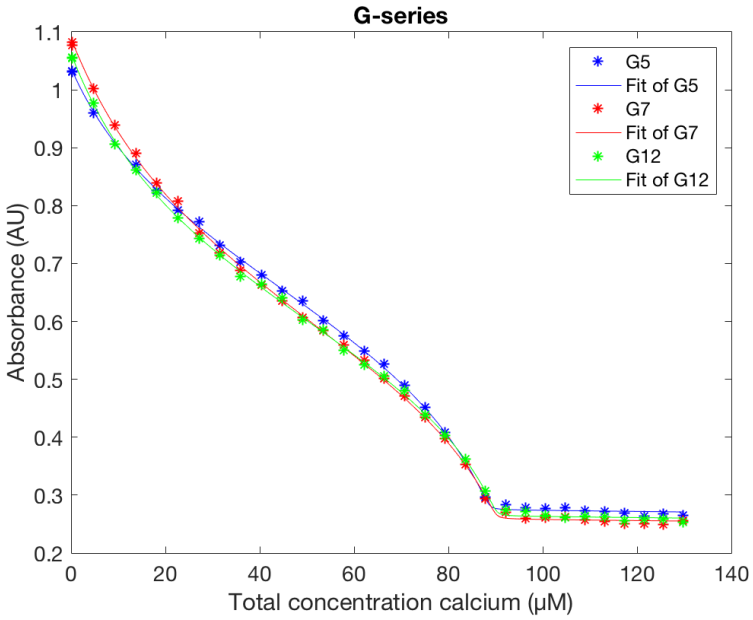

Figure 8: The absorbance of the wavelength 263 nm of the solution with both Quin-2 and the protein substrate G5, G7 or G12, plotted against the total concentration Ca^{2+} during the titration. The parameter fitting from CaLigator is showed for each substrate. The concentrations used are 33.6 µM Quin-2 and 3.0 mM stock concentration $CaCl_2$, in 2 mM TRIS buffer at pH 7.5.

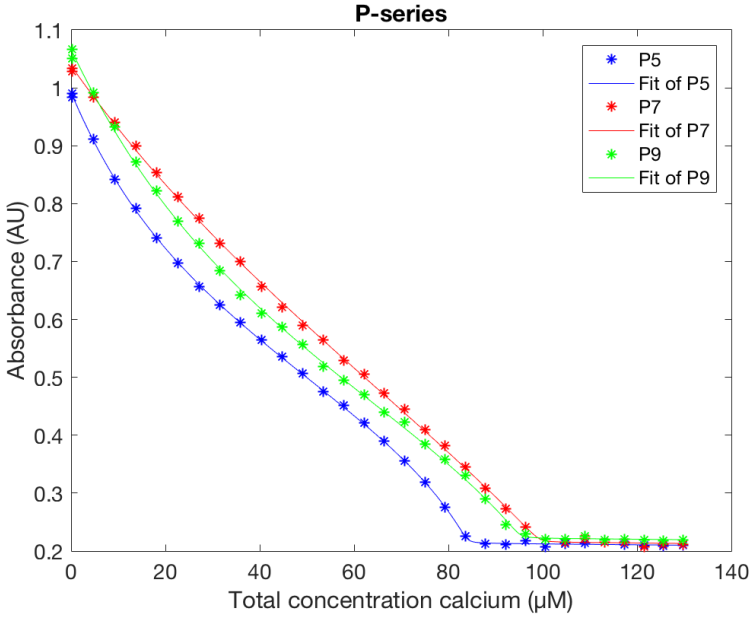

Figure 9: The absorbance of the wavelength 263 nm of the solution with both Quin-2 and the protein substrate P5, P7 or P9, plotted against the total concentration Ca^{2+} during the titration. The parameter fitting from CaLigator is showed for each substrate. The concentrations used are 33.6 µM Quin-2 and 3.0 mM stock concentration $CaCl_2$, in 2 mM TRIS buffer at pH 7.5.

In Table 2, the calculated protein concentrations and Ca^{2+} affinity constants, obtained by the software CaLigator, are presented. Both of the Quin-2 concentration estimations were used in separate fittings. The protein concentrations obtained with a Quin-2 concentration of 28.2 µM, are denoted as C_p^*, and with 33.6 µM instead, they are denoted as C_p^{**}. The Ca^{2+} affinity constants had the same values regardless of concentration used, at the decimals given in the table. The mean protein concentration of each substrate is calculated and used for later assays. They are, rounded to two digits: G5 - 27 µM, G7 - 28 µM, G12 - 28 µM, P5 - 26 µM, P7 - 32 µM, P9 - 30 µM.

Table 2: Values gained from parameter fitting with CaLigator. The protein concentration, C_p^* is obtained using the Quin-2 concentration of 28.2 µM, determined with absorbance. When instead the higher estimation of Quin-2 concentration (33.6 µM) is used, the obtained protein concentrations are given as C_p^{**}. The affinity parameters obtained in both cases differ little enough to have the same value at the decimals given in the table.

Substrate	C_p^* (µM)	C_p^{**} (µM)	$\log K_1$	$\log K_2$	$\log K_1K_2$	$-\Delta\Delta G_{min}$ (kJ/mol)
G5	24.7	29.4	7.7	8.8	16.59	9.6
G7	25.3	30.2	7.7	8.6	16.36	8.6
G12	25.3	30.2	7.6	8.8	16.39	10.5
P5	23.3	27.9	7.8	8.4	16.26	6.8
P7	29.4	35.1	8.3	8.0	16.36	1.6
P9	27.6	32.9	8.0	8.1	16.12	3.9

The product of the two Ca^{2+} binding constants, K_1K_2, is determined with higher accuracy than they are separately. This was supported by fixing K_1 respectively K_2 for values ± 0.05 from the calculated ones (since they are given in one decimal each), and by letting CaLigator estimate K_2 respectively K_1 to this. The K_1K_2 product differed well below 0.005, and hence, two decimals can be considered safe to use, given an uncertainty of ± 0.05 in K_1 and K_2 respectively.

From Figure 8, it is clear that different lenght of polyglycine linkers does not affect the Ca^{2+} affinity much. But in Figure 9 it can be seen what Table 2 tells; the different lengths of polyproline linkers affects the cooperativity greatly, with almost no positive cooperativity for P7. But the total affinity for Ca^{2+} is kept quite constant, only slightly smaller.

3.2 Dynamic Light Scattering (DLS) to Evaluate Aggregation

Is DLS a suitable method for analysing oligomeric states of S100G, and if so, in what calcium and protein concentrations do the proteins aggregate the most?

The substrate P5, at various concentrations and Ca^{2+} saturation levels, was analysed with DLS, at various times after dissolutions. The highest protein concentration, 2.5 mM, provided most consistent results and since the documentation of the instrument supports that the lower concentrations might be too low to be analysed in the instrument, only the highest concentration will be further analysed. Figure 10-13 shows the normalized autocorrelation function to the logarithmic delay time for each Ca^{2+} saturation and time point. To see the non-normalized graphs, see Appendix B, Figure 42-45.

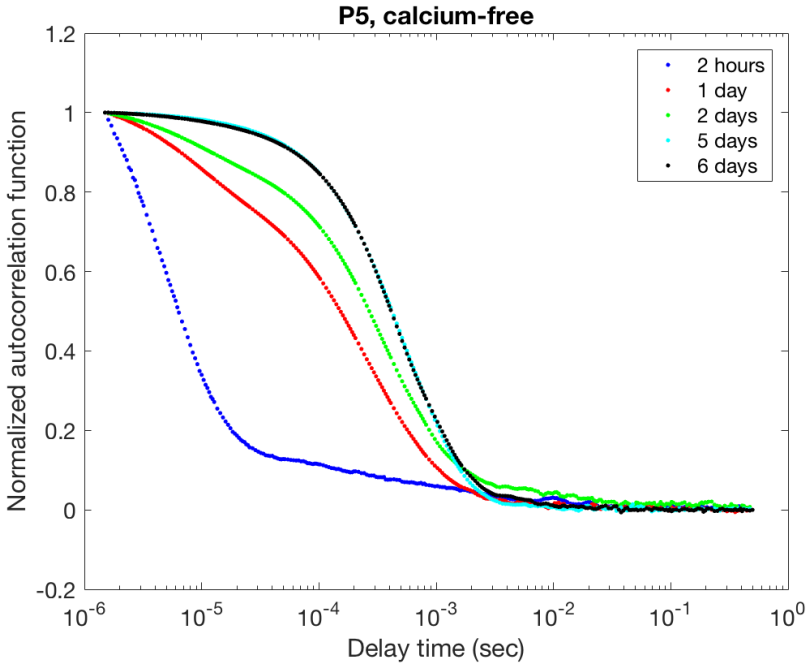

Figure 10: 2.5 mM P5 in Ca^{2+}-free buffer with 2 mM TRIS at pH 7.5, analysed using DLS. Normalized autocorrelation functions plotted against the delay time.

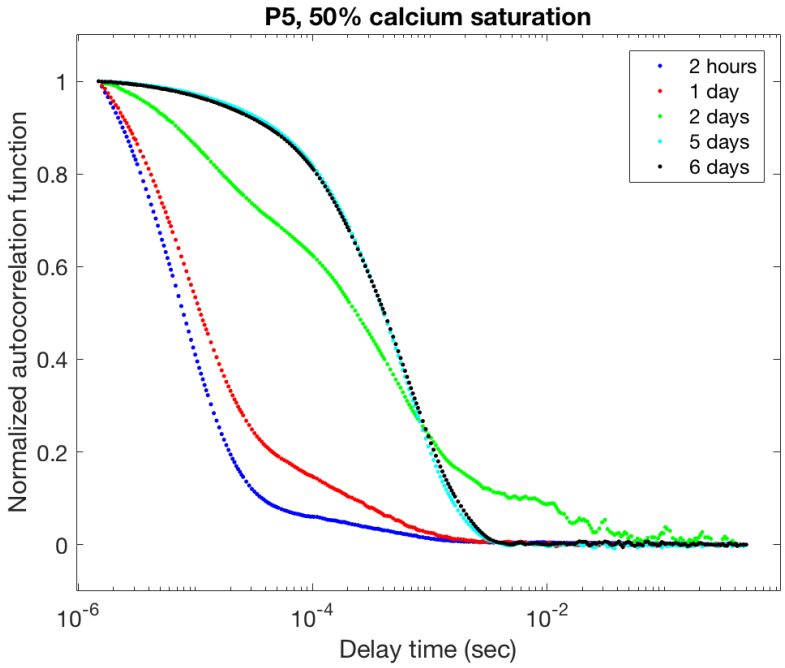

Figure 11: 2.5 mM P5 at 50% Ca^{2+} saturation with 2 mM TRIS at pH 7.5, analysed using DLS. Normalized autocorrelation functions plotted against the delay time.

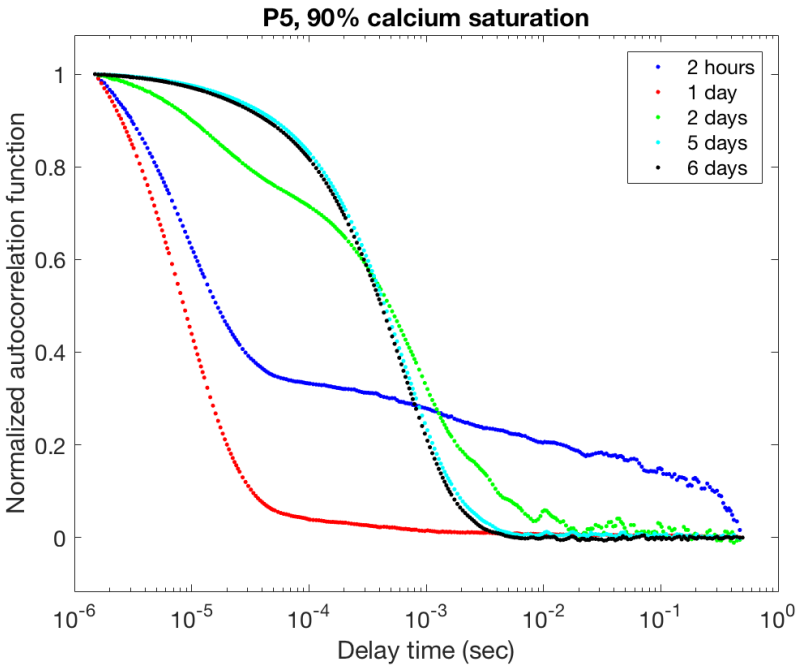

Figure 12: 2.5 mM P5 at 90% Ca^{2+} saturation with 2 mM TRIS at pH 7.5, analysed using DLS. Normalized autocorrelation functions plotted against the delay time.

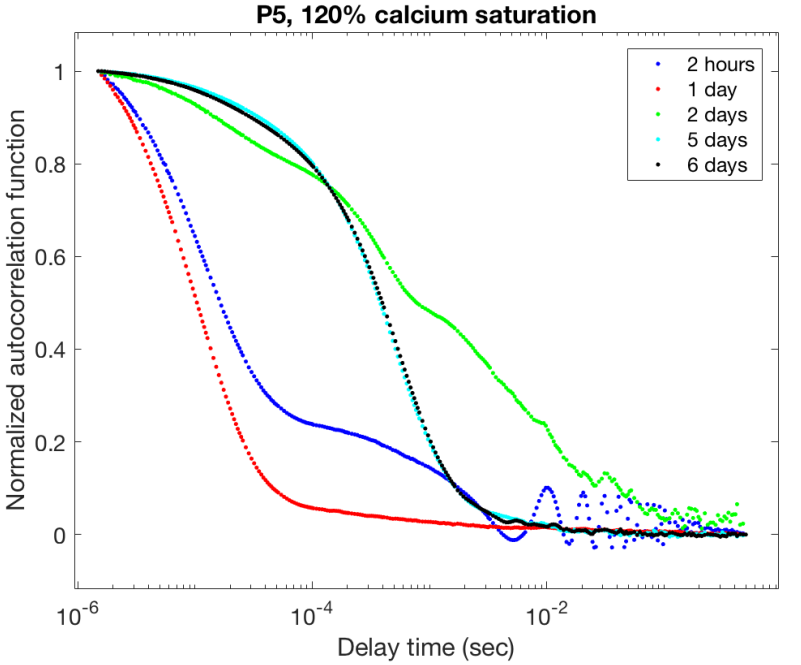

Figure 13: 2.5 mM P5 at 120% Ca^{2+} saturation with 2 mM TRIS at pH 7.5, analysed using DLS. Normalized autocorrelation functions plotted against the delay time.

For P5 in the Ca^{2+}-free buffer and at 50% Ca^{2+} saturation, a clear trend can be seen; the longer the time after protein dissolutions, the decay slope is located at longer delay

times. For the samples with 90% and 120% Ca^{2+} saturation, the data for the two hours measurement looks like it corresponds to a bit larger particles than after 1 day, and smaller than after two days. But, as argued for in the discussion section, the results indicate presence of particles which are much too large right after filtration, and thus the method seems to be a bad choice for such small particles.

3.3 Native Gel Electrophoresis to Evaluate Aggregation

Some different assays were performed using native PAGE. Firstly, screening of protein and Ca^{2+} concentrations; secondly, screening of substrates, over time; and lastly a combination assay with new P9, higher Ca^{2+} concentrations and long incubation time.

Screening of protein concentrations and Ca^{2+} saturation levels

Is it possible to detect any oligomeric states in a native PAGE, and if so, at what protein and calcium concentrations is this happening?

From the aggregation screening of the substrate P5, where protein concentrations and levels of Ca^{2+} saturation were varied, the gels in Figure 14 were obtained. Only three wells of one of the gels were used, and hence the two pictures are assembled into one figure.

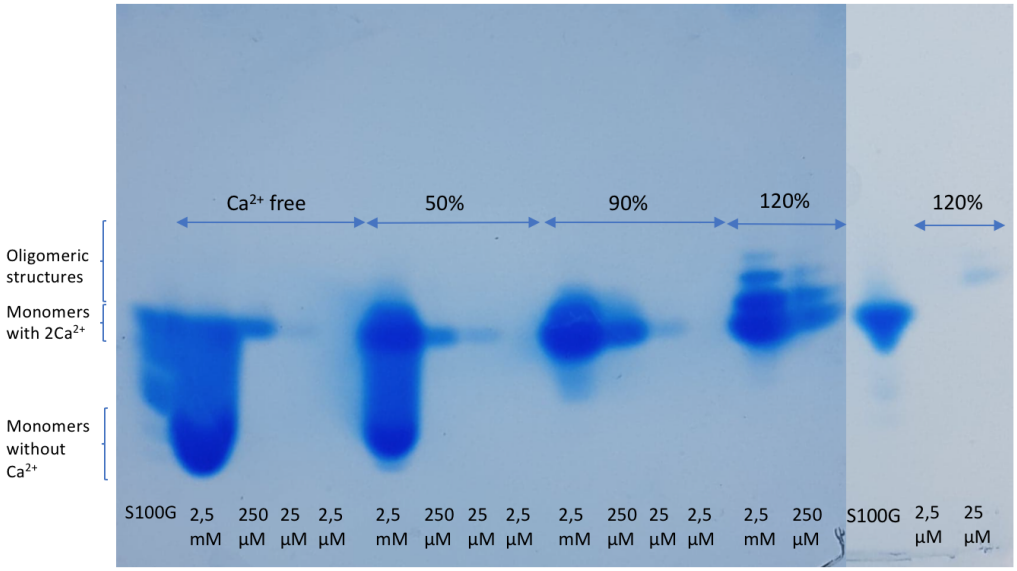

Figure 14: Native PAGE result of the aggregation screening with varying protein concentrations and Ca^{2+} saturation levels. The substrate used was P5 in 2 mM TRIS buffer, pH 7.5 and incubated for seven days at room temperature. The staining method was Instant Blue and the gels are of the type gradient 4-20% Tris-glycine gel. The different Ca^{2+} saturation series are indicated with arrows and the protein concentrations are written below each well Also, to the left it is indicated at what height different structures are located. The three wells to the right (brighter surrounding) belong to an other gel that was run in parallel, under the same conditions. The S100G references are indicated in both gels, but since the one to the left is smeared, this is likely the protein in the well next to it that have spilled over.

For the series in the Ca^{2+}-free buffer, the highest located band is at the same level as

the broadest band of the other series. The interpretation is that this level corresponds to the monomeric proteins with two bound Ca^{2+} ions. The reason this band is present in the Ca^{2+}-free series might be that the loading buffer and the running buffer are not entirely Ca^{2+}-free, causing some fraction of the proteins to bind Ca^{2+}. The broad band farthest down in the Ca^{2+}-free series (present for the highest protein concentration, 2.5 mM) corresponds to the protein without any Ca^{2+} bound, which makes it more negatively charged and thus traveling faster in the electric field. A likely reason for the smear between these bands, is that the gel might not be completely Ca^{2+}-free. When Ca^{2+}-free (apo) proteins travels through the gel, they will collect Ca^{2+} somewhere on the way, and becomes then less charged, that is, slower. The proteins with only one bound Ca^{2+} ion are also located in this region, but due to the high cooperativity, this is not a common state. It can be noted that the other protein concentrations in this series do not have this smear and only one band, which can be explained with the lower protein concentration being saturated initially of the running buffer and the loading buffer. The same explanations can be used for describing the result of the 50% Ca^{2+} saturation series.

The 90% Ca^{2+} saturation series is more compact and only a small smear downwards and upwards can be seen, indicating some unsaturated proteins initially and also some aggregation. But it is first in the 120% series that clear bands with lower mobility than monomer can be seen. For the highest protein concentration and the second highest this is most significant and up to three bands can be seen in addition to the supposed monomeric band.

Unfortunately, the S100G to the left seems to be very smeared. Likely, this is mostly proteins from the well next to it that spilled over due to too much protein in this well. The S100G to the left is located at an other gel, and can thus not be used as a reference level to the main gel.

Conclusions that can be drawn from this assay are that P5 aggregates more at Ca^{2+} levels above saturation and the aggregates were easier to see at protein concentrations between 0.25 mM to 2.5 mM.

Screening of substrates

At 80% respectively 120% Ca^{2+} saturation, how does the linker length, and type, affect the aggregation, and is it more beneficial to combine different substrates?

The aggregation screening of various substrates was done using native PAGE. The results after one, seven and fourteen days can be seen in Figure 15-17. Pictures of the gels containing samples that were two and four days old can be found in Appendix C, Figure 46-47. From the stained gels, it can be seen that the P-series aggregated mostly at 120% Ca^{2+} saturation, with P9 having the most different forms of aggregates, up to four bands. For the G-series, the difference between the two Ca^{2+} saturation levels are smaller and only up to two bands can be seen at the saturated part of the proteins (that is, highest located in each well). The smear in the gel with 120% saturation tells that the G-series was not completely saturated at this concentration, indicating the concentration determinations of the protein and/or Ca^{2+} were not perfect.

Regarding the combinations of different substrates, no additional bands besides the ones corresponding to the pure proteins could be found. Either, no aggregates were formed of more than one kind of protein, or, this is the case but it can not be detected with this assay.

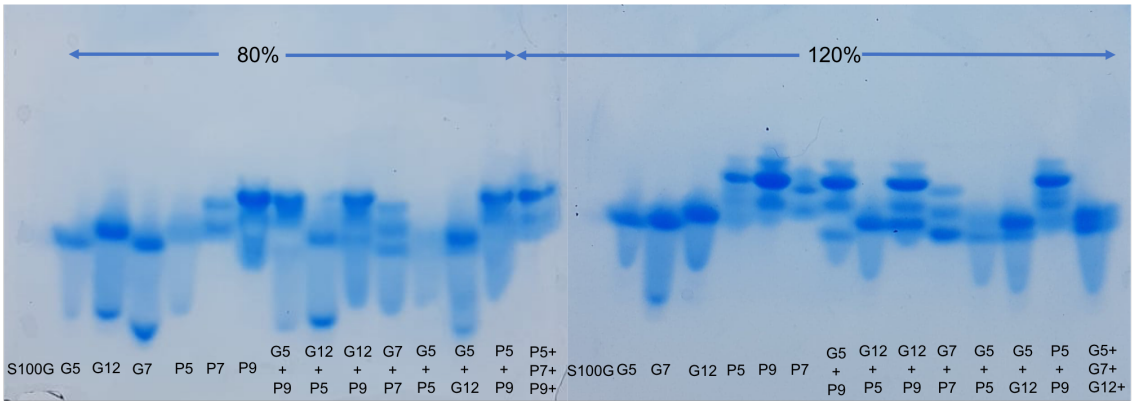

Figure 15: Native PAGE of two gels containing various substrate samples with 1 mM protein in 2 mM TRIS buffer, pH 7.5 and incubated for one day at room temperature. Which sample belongs to which well is written below each well. At each gel, a reference S100G protein is present, at surely saturated Ca^{2+} concentration, but at too low protein concentration in these gels to be seen clearly. The gels are gradient ones, 4-20% Tris-glycine and the staining method is Instant Blue.

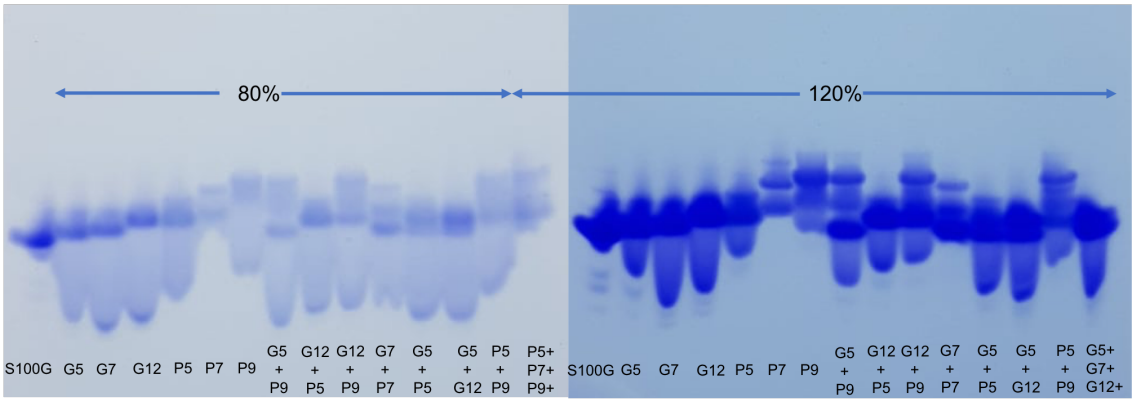

Figure 16: Native PAGE of various substrate samples with 1 mM protein in 2 mM TRIS buffer, pH 7.5 and incubated for seven days at room temperature. 14% Tris-glycine gel and the staining method is Coomassie Brilliant Blue. Here, the S100G reference was inserted with higher concentration and is clearly seen.

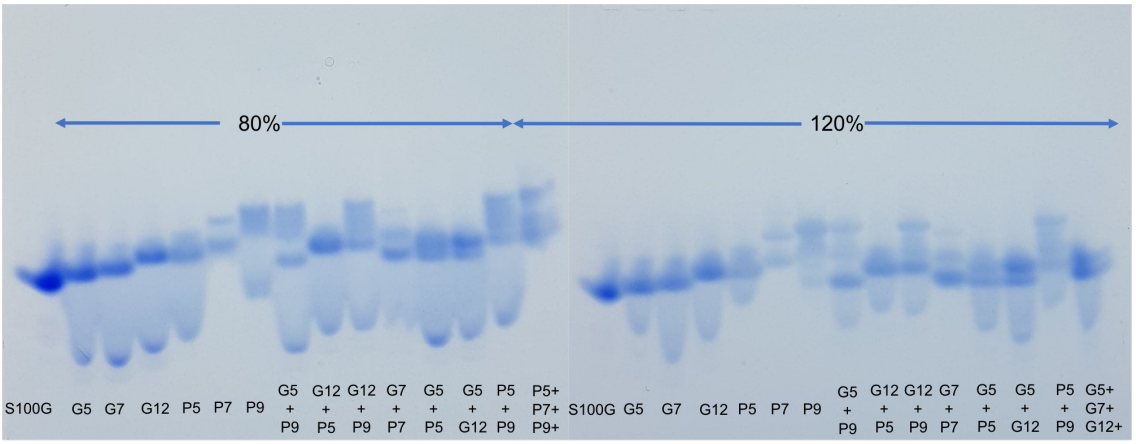

Figure 17: Native PAGE of the 1 mM protein samples in 2 mM TRIS buffer, pH 7.5 and incubated for fourteen days at room temperature. 14% Tris-glycine gel was used and the staining method is Coomassie Brilliant Blue. The S100G reference is clearly seen.

Additional PAGE for New P9, Higher Ca^{2+} Saturation Levels and a Long Time Assay

Do the newly desalted P9 aggregate in a similar way as the earlier used P9?

Is the aggregation dependent on Ca^{2+} concentration above 120% saturation?

Have the aggregates changed after 50 days of incubation?

Figure 18 shows the resulting gel from the experiment where an aliquot from the previously used P9 (called "old") can be seen in well 2, and an aliquot from the "new" P9 is in well 1. It appears that the new P9 aggregates into one or two higher oligomeric structures.

In the same gel, well 2-4, a series of Ca^{2+} saturation levels (120%, 150% and 200%) of the old P9 was included, to see if aggregation varied in this region. As can be seen, the bands look much the same in well 2-4 and thus, the conclusion is that aggregation do not appear to be significantly dependent on Ca^{2+} saturation at levels higher than 120%.

Also in this gel, some of the earlier analysed substrates and its combinations can be seen, in well 5-14. These samples have been incubated for 50 days since dissolutions, at room temperature. It can be noted that some bands have been more clear in the lowest parts of the gel. This is likely due to some degradation of the proteins. Beside from this, the G-series at 80% Ca^{2+} saturation looks similar to after 14 days of incubation (Figure 17). Focusing on the wells containing some P9, which are number 9 and 11-14, it can be seen that they do not have any band as high in the gel as P9 in well 2-4. Hence, some of the larger aggregates seems to have broken down after this time.

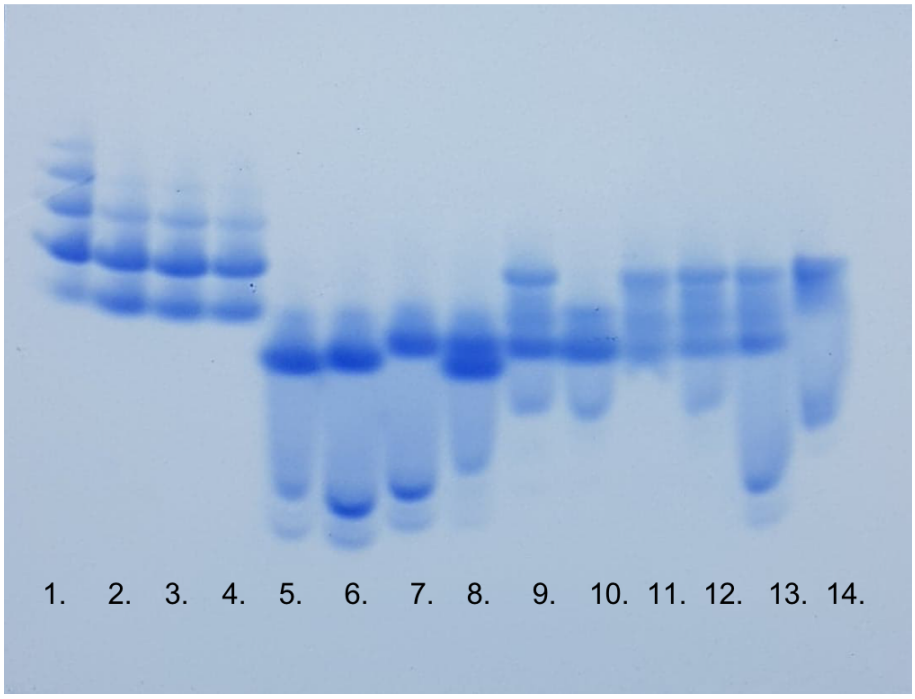

Figure 18: A native PAGE where the wells are numbered 1-14. All protein concentrations is 1 mM and pH is 7.5. The P9 substrates in the four wells to the left were incubated in one hour at room temperature and the other wells contained some of the earlier substrates that now had been incubated for 50 days at room temperature. From left to right, with the Ca^{2+} saturation in brackets: 1. new P9 (120%), 2. old P9 (120%), 3. old P9 (150%), 4. old P9 (200%), 5. G5 (80%), 6. G7 (80%), 7. G12 (80%), 8. G5+G7+G12 (120%), 9. G12+P9 (120%), 10. P5 (120%), 11. P5+P7+P9 (120%), 12. P5+P9 (120%), 13. G12+P9 (80%), 14. P9 (80%).

3.4 MDS for Size Determination of Aggregates

Can MDS be used to support the native PAGE results?

What sizes have different oligomers of P5, P7 respectively P9?

About 24 hours after dissolutions in 2 mM HEPES buffer, pH 7.5, the protein solutions were analysed using MDS and the output values calculated by the instrument are presented in Table 3. A mean value of each triplicate is calculated for each parameter.

Table 3: The results from MDS with the P-series after one day at room temperature and S100G without inserted linker. The instrument reported a calculated hydrodynamic radius, R_H, and a protein concentration, C_P. For each protein solution, three runs where made and a mean value was calculated. The mean value of the hydrodynamic radius of P5 at 80% Ca^{2+} saturation is calculated from only run one and two (hence the asterisk). This was done since the raw data of the third run did not have a smooth curve in diffusion proportion, as the others had. In an other assay, but under the same conditions, also P7 dissolved in Ca^{2+}-free buffer were analysed, presented last in the table.

1 day	**R_H (nm)**				**C_P (µg/ml)**			
Replicate	**1**	**2**	**3**	**Mean**	**1**	**2**	**3**	**Mean**
S100G	1.17	1.11	1.25	1.18	455	442	460	452
P5 Ca^{2+}-free	1.14	1.14	1.22	1.17	589	567	580	579
P5 80% saturation	1.30	1.39	1.95	1.35*	639	548	514	567
P5 120% saturation	1.41	1.47	1.56	1.48	636	641	646	641
P9 Ca^{2+}-free	1.27	1.35	1.39	1.34	598	479	498	525
P9 80% saturation	1.55	1.52	1.59	1.55	478	445	457	460
P9 120% saturation	1.83	1.86	1.97	1.89	466	425	426	439
P7 Ca^{2+}-free	1.52	1.62	1.58	1.57	164	193	232	196

From these results, it can be said that there is a strong indication that with 80% or 120% Ca^{2+} saturation, the particles are larger than they are under low-saturated conditions. The natural explanation to this is that P5 and P9 both aggregates more with 80% and 120% Ca^{2+} saturation. This is in line with the native PAGE assay that compared different saturation levels. It shall be noted that it is the mean hydrodynamic radius that is given by this method, and from this, the distribution of the particle sizes can not be concluded.

The result from the measurements performed one week after the time the protein solutions were made are presented in Table 4. The mean hydrodynamic radius have not changed much over these days.

Table 4: By the instrument calculated hydrodynamic radius, R_H, and protein concentration, C_P for the protein solutions after seven days at room temperature.

7 days	**R_H (nm)**				**C_P (µg/ml)**			
Replicate	**1**	**2**	**3**	**Mean**	**1**	**2**	**3**	**Mean**
P5 Ca^{2+}-free	1.31	1.28	1.25	1.28	787	779	802	789
P5 80% saturation	1.37	1.43	1.51	1.44	314	311	317	315
P5 120% saturation	1.29	1.43	1.50	1.41	405	415	424	415
P9 80% saturation	1.60	1.53	1.59	1.57	502	431	344	426
P9 120% saturation	1.74	1.97	1.99	1.90	262	257	271	263

Measurements with MDS were also performed of the fractions containing separated aggregates from the SEC eluate. The results can be viewed in Table 5 for substrates that were analysed using SEC a few minutes after dissolutions and in Table 6 for substrates

that had been incubated for seven days. The MDS instrument did not work perfectly for measurements at time point 0 hours, which is why not all peaks were analysed.

Table 5: Hydrodynamic radii and protein concentrations measured with MDS of SEC eluate. Mean values are calculated for each triplicate. The substrates were analysed as soon as possible after dissolutions. The assay was aborted at P7, fourth peak, since the instrument only gave errors.

0 hours, after SEC	**R_H (nm)**				**C_P (µg/ml)**			
Replicate	**1**	**2**	**3**	**Mean**	**1**	**2**	**3**	**Mean**
P5, peak 1	3.16	2.95	2.82	2.98	12.9	13.0	13.8	13.2
P5, peak 2	2.42	2.53	2.58	2.51	95.9	99.2	95.3	96.8
P5, peak 3	1.80	2.07	1.75	1.87	113	111	119	114
P7, peak 1	9.91	2.91	2.93	5.25	2.05	3.38	1.60	2.34
P7, peak 2	2.48	3.49	5.61	3.86	5.20	6.94	8.77	6.97
P7, peak 4	-	-	-	-	-	-	-	-

Table 6: Hydrodynamic radii and protein concentrations measured with MDS of eluate from SEC. Mean values are calculated for each triplicate. The substrates were analysed after incubation in seven days since dissolutions.

7 days, after SEC	**R_H (nm)**				**C_P (µg/ml)**			
Replicate	**1**	**2**	**3**	**Mean**	**1**	**2**	**3**	**Mean**
P5, peak 1	2.83	2.67	3.01	2.84	1.14	1.01	1.23	1.13
P5, peak 2	2.60	2.63	2.65	2.63	142	136	142	140
P5, peak 3	1.69	1.75	1.79	1.74	206	205	212	208
P7, peak 1	2.93	2.66	3.47	3.02	1.96	2.00	1.84	1.94
P7, peak 2	3.13	3.18	3.26	3.19	69.8	70.3	65.0	68.4
P7, peak 3	2.50	2.67	3.16	2.78	341	353	350	348
P7, peak 4	1.81	1.80	1.85	1.82	232	231	225	229
P9, peak 1	3.24	3.73	3.22	3.40	5.32	5.00	5.32	5.21
P9, peak 2	3.37	3.48	3.56	3.47	19.9	20.7	20.8	20.5
P9, peak 3	2.67	2.78	2.77	2.74	101	98.2	97.6	99.0
P9, peak 4	1.82	1.92	1.89	1.88	13.0	13.2	12.6	12.9

It should be noted that the protein concentrations differs a lot between the different peak-fractions. This is due to the various height of the peaks from the SEC, which can be seen in Figure 23-25. Unfortunately, the concentration of the largest aggregates is too low for the instrument to analyse them accurately (limit of detection is 10 µM). It is therefore hard to tell how reliable these measurements are, but given the non-consistent results, as for instance of P7, 0 hours, first peak, not much weight are recommended to put to the results of low concentrations.

3.5 SEC for the Evaluation of Aggregation

3.5.1 General Treatment of the Raw Data

The absorbance of various wavelengths and the conductivity of the eluate were measured. For each run, these measurements have been plotted to the eluate volume since start of run. An example can be seen in Figure 19. The conductivity is constant except for a well just before 10 mL, arising from a smaller ionic strength of the buffer the proteins were dissolved in, compared to the running buffer. In the region 15-20 mL, two absorbance peaks for wavelengths 260 nm and 280 nm are present, but the 214 nm absorbance is very low here and hence the analyse tells that this is not proteins. Similar features can be seen in the other runs. To easier analyse the peaks arising from proteins, the following graphs are restricted to volumes between 7 and 16 mL and the conductivity is not shown since it is constant in this region and thus not useful for this analyse.

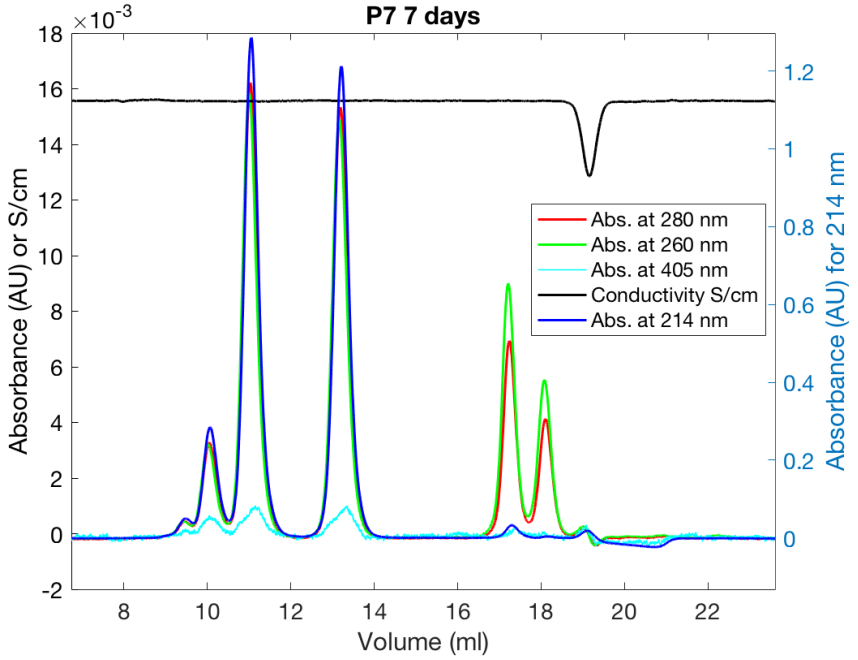

Figure 19: An example of the raw data from the SEC measurements. The substrate was 1 mM P7 at 120% Ca^{2+} saturation, 2 mM HEPES buffer and pH 7.5, incubated at room temperature for seven days. The absorbance of the wavelength 214 nm is much higher than the other and thus, it is shown at its own axis. Also note that the left y-axis unit of the conductivity is S/cm.

3.5.2 Stability of the Aggregates

Are the aggregates fast forming and reforming, and how stable are they diluted?

To evaluate the stability of the aggregates, fractions of the peaks were analysed using SEC again. For this purpose, P5 at pH 7.5 after seven days of incubation at room temperature were used. The result is showed in Figure 20. The absorbances are scaled to make it possible to analyse all three runs in the same plot. It can clearly be seen that the aggregates mostly remain intact in the time between the runs, 30 minutes for first peak and 60 minutes for the second peak. But some restructuring occurs since smaller peaks are visible at the same

positions as the other peaks. This phenomena is more pronounced for the first peak, and since the time between the runs is shorter for this, it can be said that the largest aggregates seem to be less stable than the aggregates of the second peak.

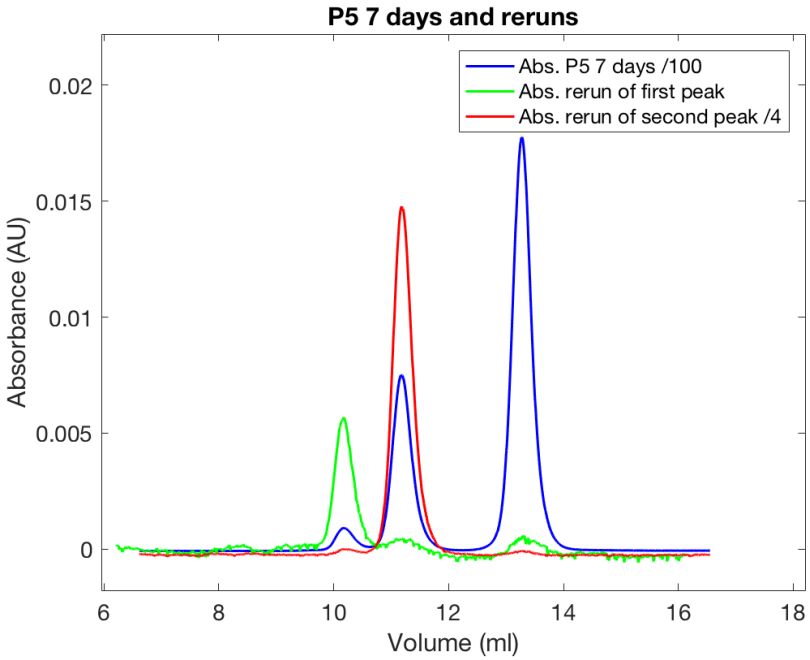

Figure 20: The absorbance of the wavelength 214 nm of P5, incubated for seven days at room temperature, and of reruns of the same with a fraction of each peak corresponding to the larger particles. To analyse the graphs in the same plot, the absorbance of the whole sample (blue) is divided by 100 and the absorbance of the rerun of the middle peak (red) is divided by 4.

Figure 21 and 22 show the results of the SEC assay where no $CaCl_2$ was added to the running buffer. The absorbance of 214 nm is plotted to the eluate volume since the time point of injection. In Figure 21, the absorbance is normalized with respect to the second peak (smallest particles), to enable comparison of height for of the first peak. Se Appendix D, Figure 48, for non-normalized graphs. It can be seen that all three substrates have bellies between the two peaks.

To see how stable the peaks are, fractions of P9 eluate was run one more time. This can be viewed in Figure 22. Interestingly, fractions from both peaks gave similar result; only one peak, at the volume corresponding to the smallest particles. Also, nothing of the belly can be seen. This shows that the aggregates are not very stable under the conditions of the running buffer, at least not on the time scale of about 30 minutes.

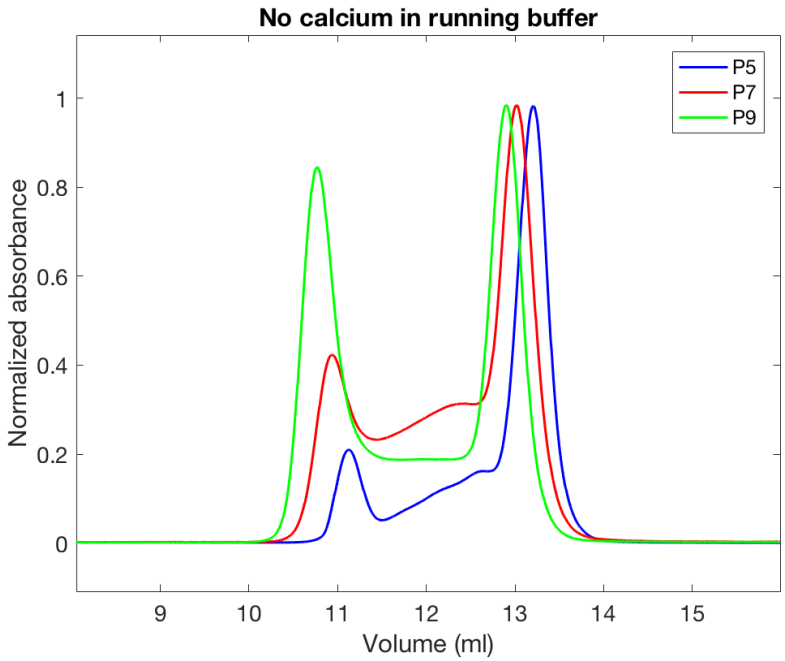

Figure 21: The absorbance of the wavelength 214 nm, measured of P5, P7 and P9 eluates. The absorbances are normalized with respect to the second peak. The substrates were analysed using SEC directly after dissolving to 1 mM protein in 2 mM TRIS buffer, pH 7.5. No Ca^{2+} were added to the running buffer, but the samples had a saturation of 120% before injection.

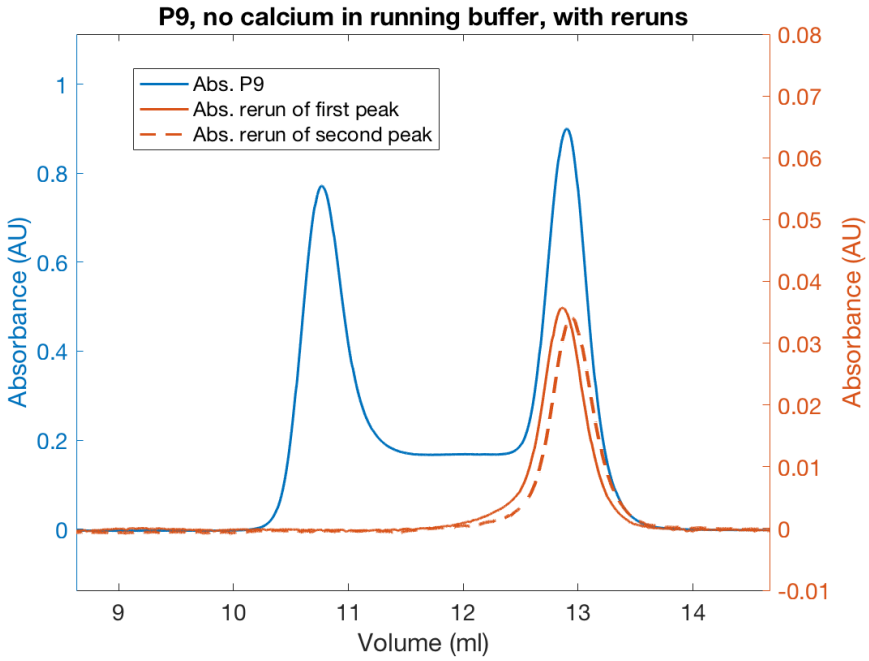

Figure 22: The same absorbance of P9 as in Figure 21. Also plotted are the reruns of the first and second peak fractions of the P9 eluate, ran 30-60 minutes later. Due to lower signal, they are plotted to another y-axis (red).

3.5.3 Time Dependence of Aggregation

How do the aggregates change with the time since dissolutions?

For P5, P7 and P9, at pH 7.5 and at room temperature, samples were analysed with SEC after various times from the point the proteins were dissolved in buffer. The results are displayed in Figure 23-25. Each plot consist of the substrates measured at the times: a few minutes after dissolution, after two respectively seven days. The absorbances are normalised with respect to the last peak, enabling qualitative comparison for the other peaks. Se Appendix D for non-normalized graphs, Figure 49-51.

To compare the substrates with each other, and to a only slightly mutated version of S100G (P43M), Figure 26 was generated and shows all their spectra for 214 nm light and after 2 days of incubation at room temperature (except for S100G that was ran a few minutes after dissolving since it is a reference of monomeric structure). The absorbances are normalized with respect to the last peak (non-normalised spectra can be seen in Appendix D, Figure 52). From this, it can be seen that the high peak of S100G is located at highest eluate volume, corresponding to smallest particles. The last peaks of the P-series proteins are only slightly shifted to the left, which tells that these particles are monomeric forms of their respectively proteins. The peaks of P5 are a little more delayed than P7, which in turn has peaks somewhat more shifted to the right than P9 has. This is in line with the additional prolines that differs between the substrates in size. But besides the shifts, the peaks of the P-series seem to arise due to aggregates that contain equally many proteins. It is also clear, from Figure 23-25 that P5 has one less significant peak than P7 and P9, and that its larger particles are less common compared to P7 and P9.

In the captions to Figure 23-25, some differences between the substrates are highlighted and interpretations stated. But it is also important to consider the similarities of the spectra. All of them keeps their shapes relatively well from a few minutes to one week. No additional peaks arise and none disappear, they only change in height.

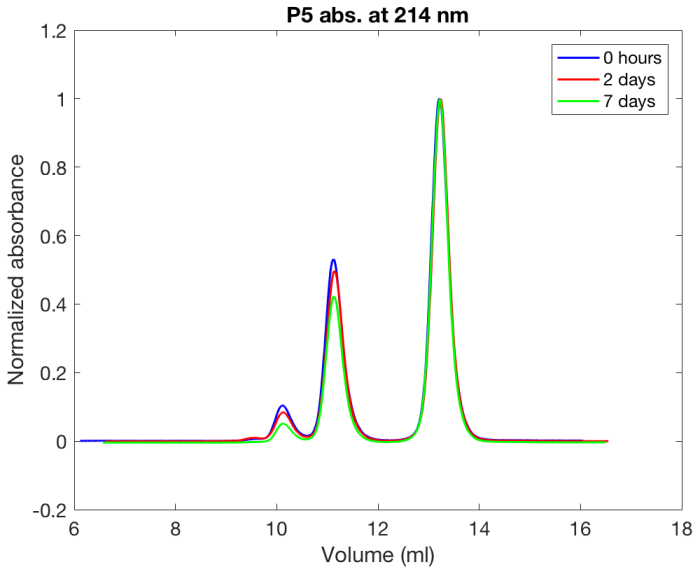

Figure 23: The absorbance of the wavelength 214 nm of P5 eluate from SEC. The protein was incubated at the concentration 1 mM, in 2 mM HEPES buffer, pH 7.5, at 120% Ca^{2+} saturation and at room temperature. The measurements were done at the time points: directly after dissolution, after two days respectively seven days. Each absorbance series is normalized with respect to the last peak. An interpretation is that larger aggregates become less and less prevalent the longer the time after dissolution.

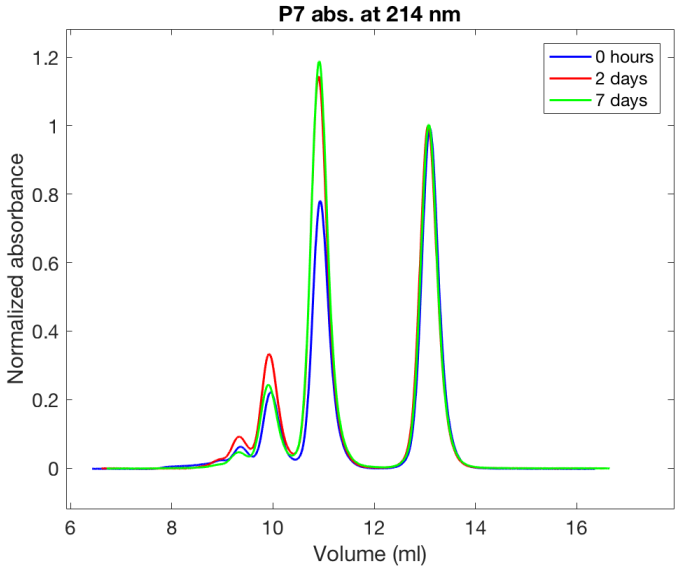

Figure 24: The absorbance of the wavelength 214 nm of P7 eluate from SEC. The protein was incubated at the concentration 1 mM, in 2 mM HEPES buffer, pH 7.5, at 120% Ca^{2+} saturation and at room temperature. The measurements were done at the time points: directly after dissolution, after two days respectively seven days. Each absorbance series is normalized with respect to the last peak. For P7, it seems that larger aggregates are most prevalent after 2 days, and than less so after 7 days. But the second smallest particles increase with time.

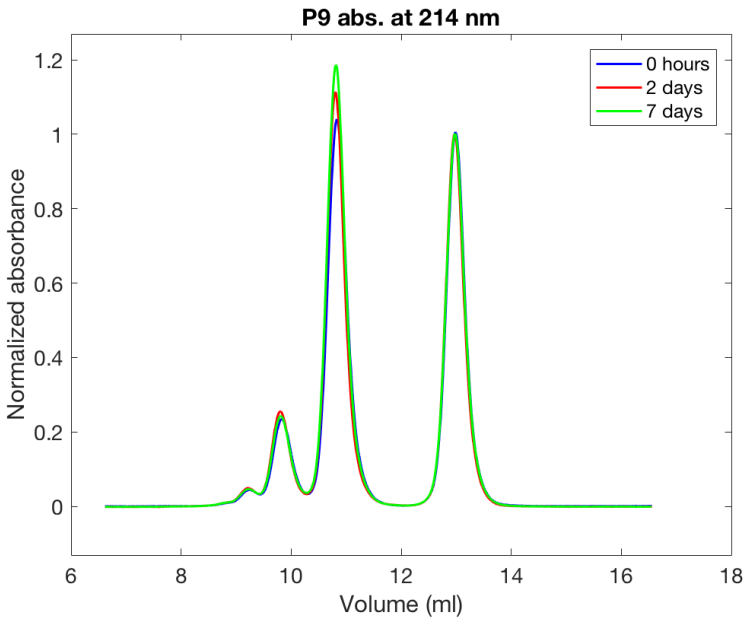

Figure 25: The absorbance of the wavelength 214 nm of P9 eluate from SEC. The protein was incubated at the concentration 1 mM, in 2 mM HEPES buffer, pH 7.5, at 120% Ca^{2+} saturation and at room temperature. The measurements were done at the time points: directly after dissolution, after two days respectively seven days. Each absorbance series is normalized with respect to the last peak. For this substrate, the second smallest aggregates increases slightly in occurrence with time, but the other seems to remain quite constant.

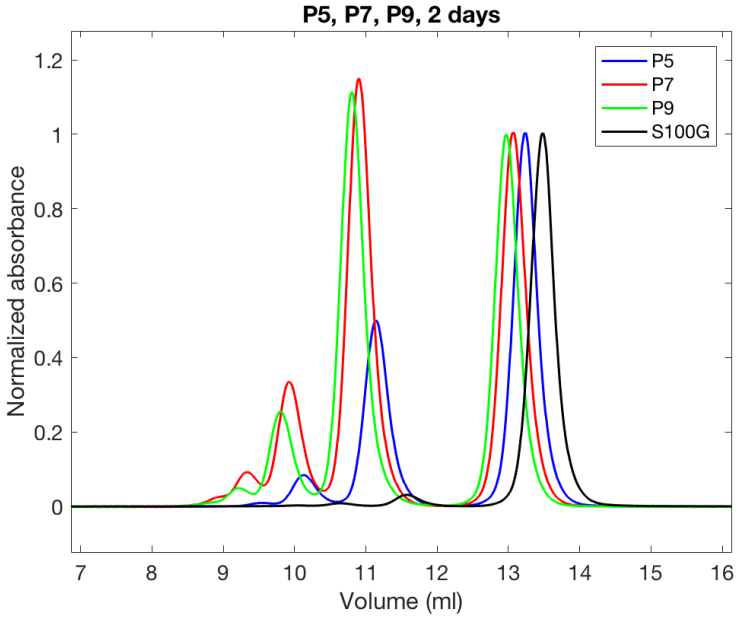

Figure 26: Comparison of absorbance (214 nm light) of P5, P7 and P9. Each absorbance series is normalized with respect to the last peak. All of the proteins were incubated for 2 days at the concentration 1 mM, in 2 mM HEPES buffer, pH 7.5, at 120% Ca^{2+} saturation and at room temperature.

3.5.4 Temperature Dependence of Aggregation

How does the temperature affect what aggregates are formed?

Only P9 was used for this evaluation, at pH 7.5. Four samples were incubated for one day at different temperatures and analysed with SEC. Figure 27 shows the absorbance at wavelength 214 nm, normalized with respect to the peak of the smallest particles. In Appendix D, Figure 53, a plot of measurements performed in a similar way, one hour after the proteins were dissolved in buffer, can be found. But these results are very similar to the ones after one day. The non-normalized graphs are located in Appendix D, Figure 54-55. The largest aggregates are less favoured at 70°C than at the other temperatures and at 50°C, the second smallest particles seems to thrive the most. The graph corresponding to the sample at 4°C is almost totally covered by the one for 37°C and is thus hard to distinguish. To get a closer look of this, see figure 28. From this, it is likely to say that the aggregation has a very small dependence for temperatures between 4°C and 37°C, but more studies needs in this interval to tell with higher certainty.

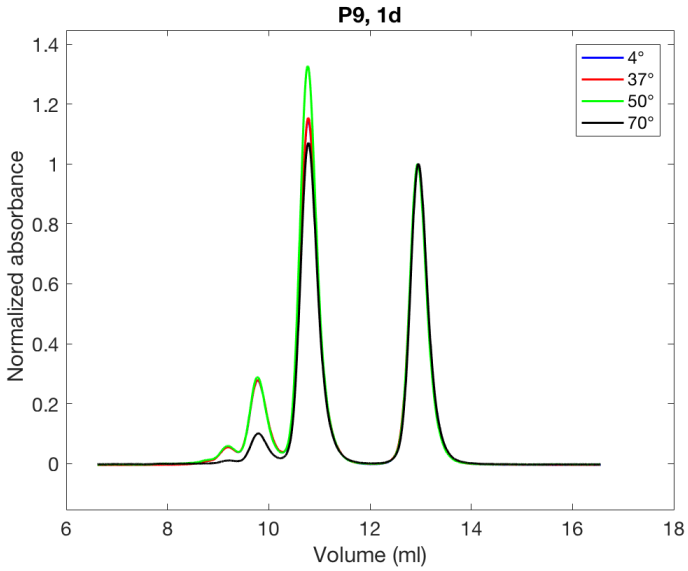

Figure 27: Absorbance (214 nm light) of eluate of P9 that had been incubated for one day at various temperatures. Protein concentration was 1 mM, in 2 mM HEPES buffer, pH 7.5 and at 120% Ca^{2+} saturation. The absorbance are normalized with respect to the last peak.

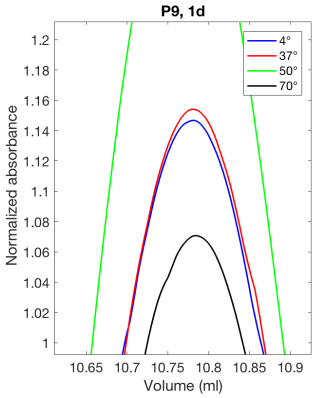

Figure 28: A closer look at the highest peak in Figure 27. Here it can be seen that the curve corresponding to 4°C (red) and to 37°C (blue) overlap closely.

3.5.5 pH and Ionic Strength Dependence of Aggregation

How does the pH affects what aggregates are formed?

How does the addition of 0.15 M NaCl affect what aggregates are formed?

The absorbance spectra of the wavelength 214 nm of the SEC eluate of P9 are displayed in Figure 29-32. They are all normalized with respect to the peak corresponding to the smallest particles, allowing for comparison of the heights of the other peaks. This is also the case for Figure 33, which compares the different pH after one hour incubation at 50°C.

Firstly, it should be noted that all spectra in these plots are very similar to shape, so it looks like the aggregation is not greatly dependent of either pH or NaCl concentration in these ranges. All four pH have very similar shapes, which can be seen in Figure 33, only for pH 9, it looks like the largest aggregates are a bit less common, even though it is not a big difference.

However, there are some differences of significance in Figure 29-32. It can be seen that the highest peak, corresponding to the second smallest particles, is higher after one day than after one hour, and this is the case both when varying pH and NaCl concentrations. Thus, the dimeric state seems to be more energetically beneficial than higher oligomers. This tendency is more pronounced in presence of 0.15 M NaCl, suggesting that the higher ionic strength makes the dimers more stale, or that it lowers the energy barriers between the oligomeric states, speeding up the formation of dimers.

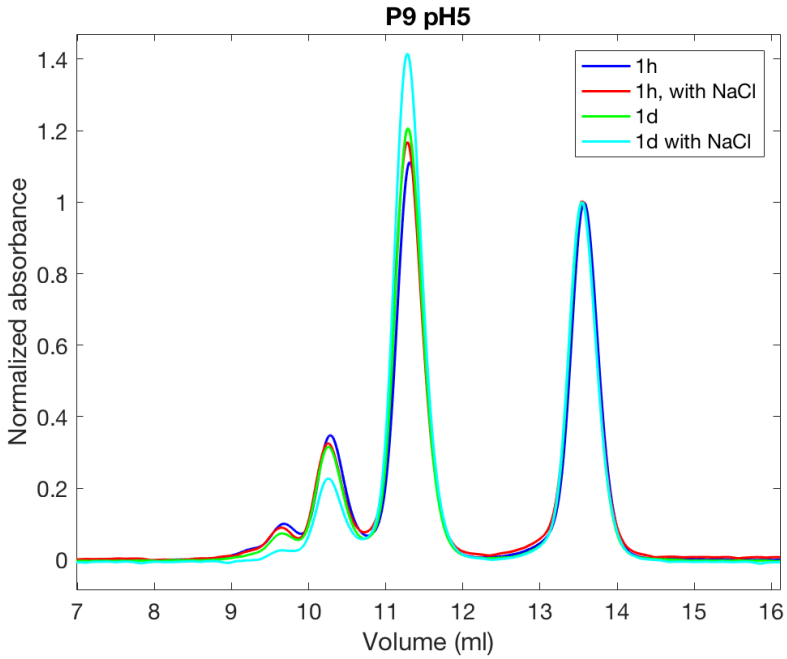

Figure 29: Absorbance (214 nm light) of eluate of P9 at pH 5, with and without 0.15 M NaCl. Incubated for one hour respectively for one day at protein concentration 1 mM, in 2 mM MES monohydrate buffer, at 120% Ca^{2+} saturation and at 50°C. The absorbance is normalized with respect to the last peak.

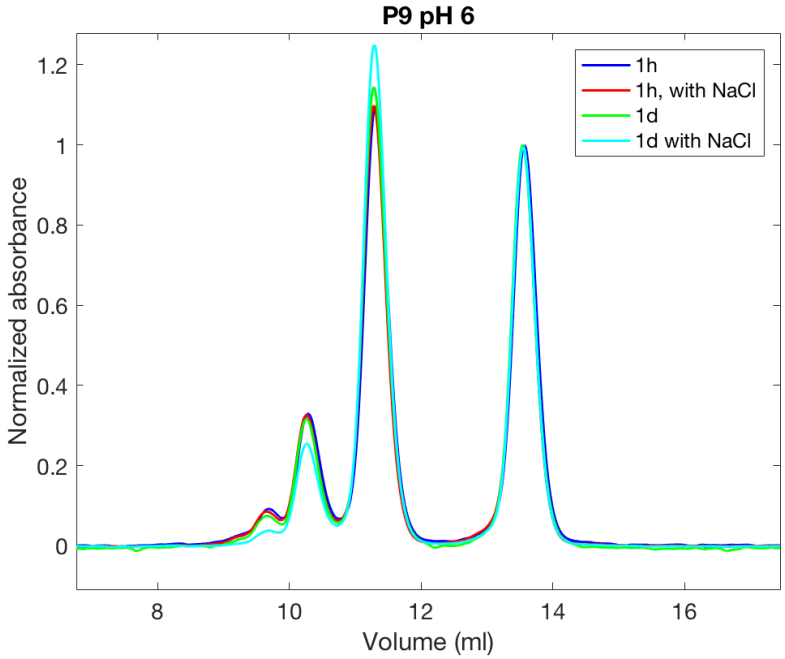

Figure 30: Absorbance (214 nm light) of eluate of P9 at pH 6, with and without 0.15 M NaCl. Incubated for one hour respectively for one day at protein concentration 1 mM, in 2 mM MES monohydrate buffer, at 120% Ca^{2+} saturation and at 50°C. The absorbance is normalized with respect to the last peak.

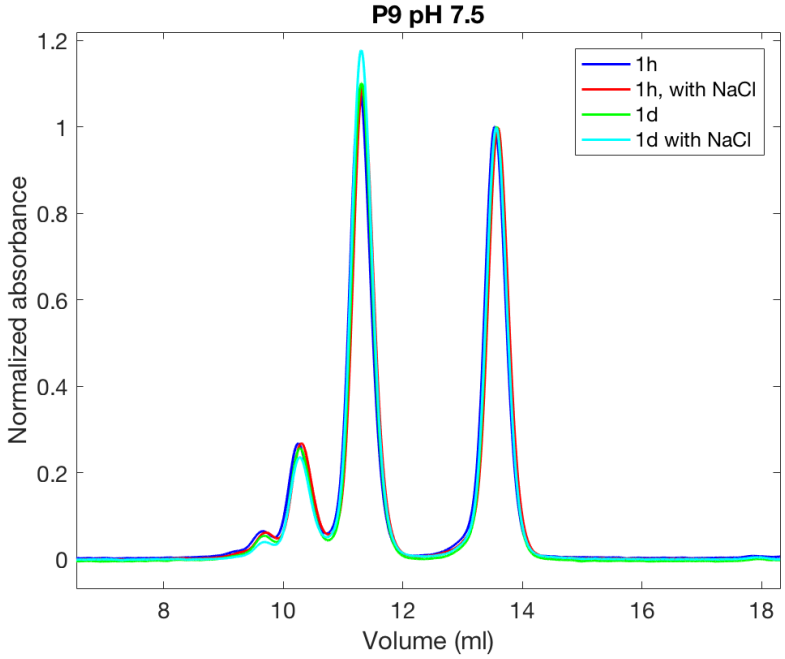

Figure 31: Absorbance (214 nm light) of eluate of P9 at pH 7.5, with and without 0.15 M NaCl. Incubated for one hour respectively for one day at protein concentration 1 mM, in 2 mM HEPES buffer at 120% Ca^{2+} saturation and at 50°C. The absorbance is normalized with respect to the last peak.

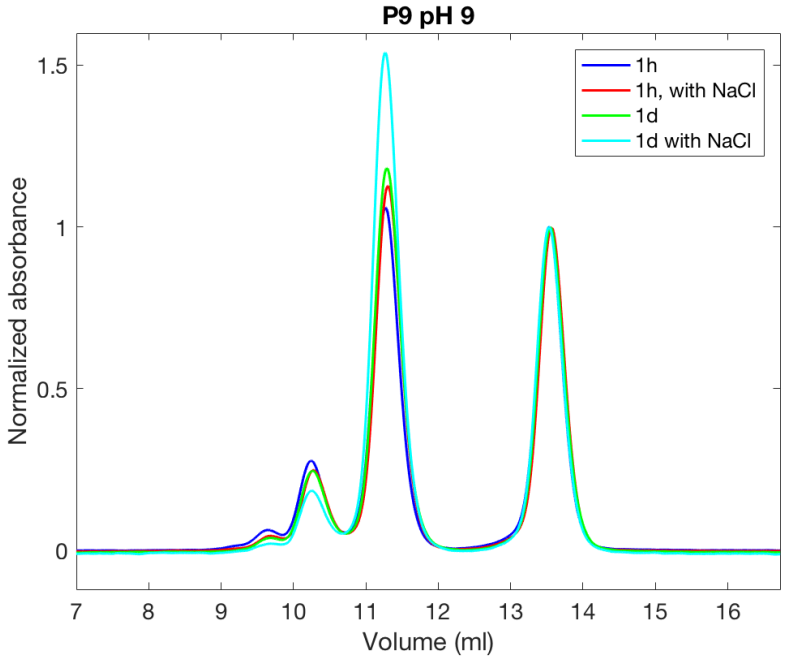

Figure 32: Absorbance (214 nm light) of eluate of P9 at pH 9, with and without 0.15 M NaCl. Incubated for one hour respectively for one day at protein concentration 1 mM, in 2 mM Tricine buffer at 120% Ca^{2+} saturation and at 50°C. The absorbance is normalized with respect to the last peak.

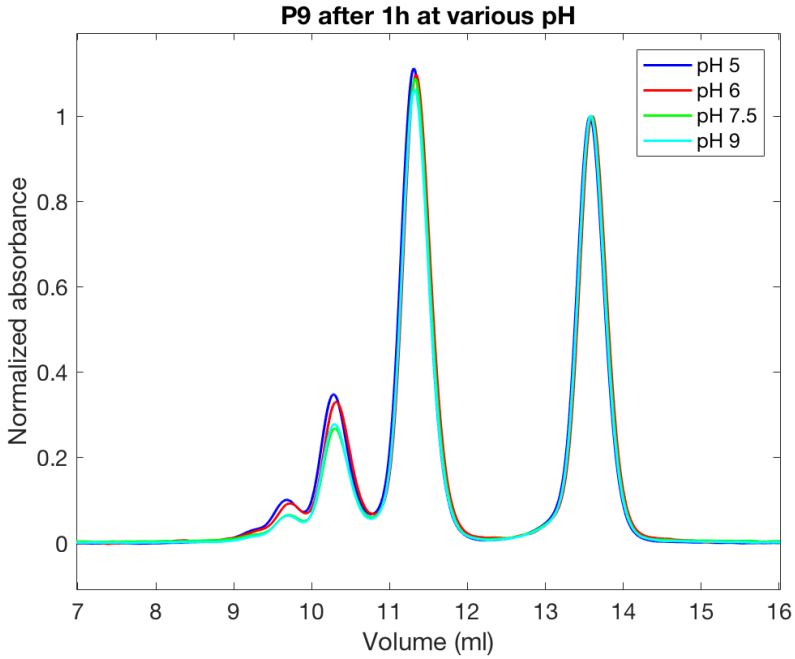

Figure 33: Comparison of P9 at all four different pH, incubated for one hour at 50°C. The absorbance is normalized with respect to the last peak.

From the assay evaluating the method of having different pH in sample and column, the

absorbance spectrum of 214 nm light is shown in Figure 34. To compare with, the spectrum of an equal sample run in the normal running buffer (same data as showed in Figure 29, blue curve). The graphs are normalized with respect to the height of the last peak. The non-normalized spectra can be found in Appendix D, Figure 56-60. Overall, the graphs have similar shapes, with peaks at the same volumes. But the column in pH 5 provides fewer aggregates of larger kinds and a small additional peak can be seen at 8 ml. The void volume of this column is 7.2 ml. This indicates that the sample contained (very little of) something in the size of the upper limit of the separation range of the column, which is 3-70 kDa for a globular protein.

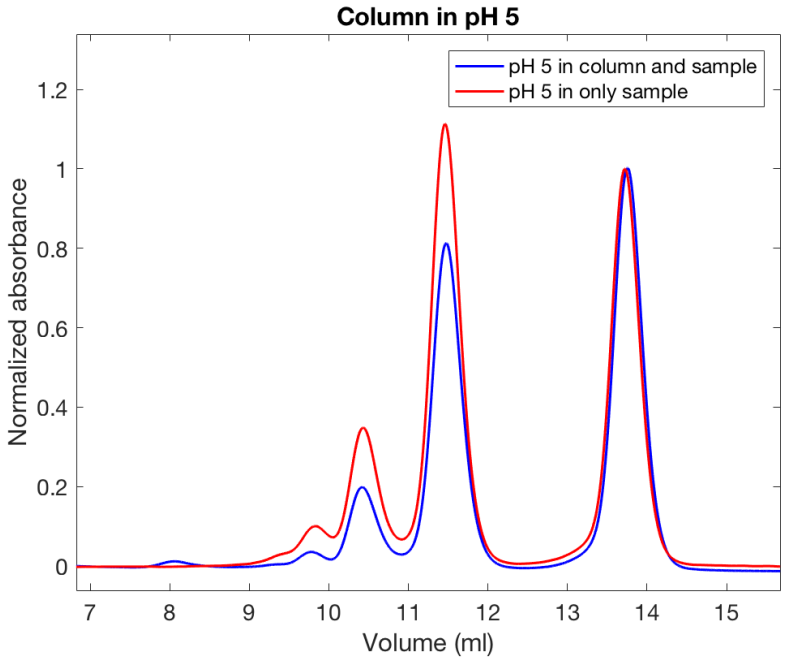

Figure 34: Absorbance (214 nm light) of eluate of P9, incubated for one hour at protein concentration 1mM, in 2 mM MES monohydrate buffer, pH 5, at 120% Ca^{2+} saturation and at 50°C. The blue graph corresponds to the run where the running buffer had pH 5 and the red to the normal running buffer of pH 7.5. The absorbance is normalized with respect to the last peak.

3.5.6 Comparison of Newly Desalted P9 with Earlier Used P9 Aliquots

Do the newly desalted P9 aggregate in a similar way as the earlier used P9?

As a complement to the native PAGE which evaluated this question, also SEC was used. The analysis was performed with three different columns, in an attempt to separate also the largest particles present. The absorbance of 214 nm light, of eluate from the Superdex 75 Increase, is shown in Figure 35, both for the new and old P9. The absorbance is normalized with respect to the second last peak and a non-normalized plot can be found in Appendix D, Figure 62. The large peak at about 7.5 ml eluate corresponds to the void volume of the column. Between this peak and the other peaks that correlate with the old P9, a belly can be seen and also some very small additional peak(s).

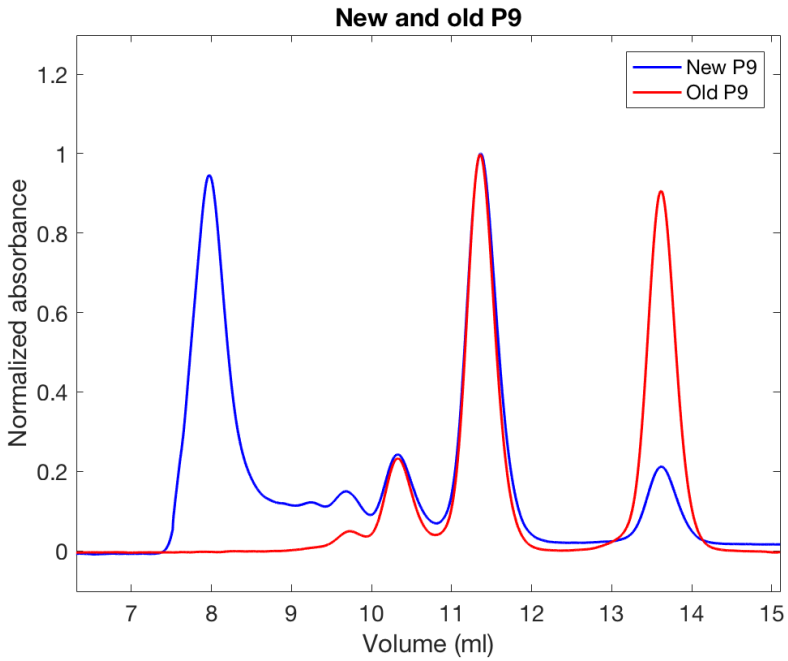

Figure 35: Absorbance (214 nm light) of eluate of the new and old P9, separated in Superdex 75 Increase. The proteins were incubated for one hour at protein concentration 1mM, in 2 mM HEPES buffer, pH 7.5, at 120% Ca^{2+} saturation and at room temperature. The absorbance is normalized with respect to the second last peak.

When a column of the type Superdex 200 Increase was used for the new P9 (otherwise same conditions as above), the resulting absorbance of the eluate can be seen in Figure 36. Also showed are the runs made with collected fractions at the eluate volume of 7.5-8 ml (corresponding to the large particles in the void volume) respectively at 10-10.5 ml (in the middle of the belly). Since the absorbance of the re-run of the belly fraction is much weaker, it is plotted to another y-axis (red). The re-run peak of the largest particles is located about 0.5 ml later than earlier and the other re-run has kept its location similar.

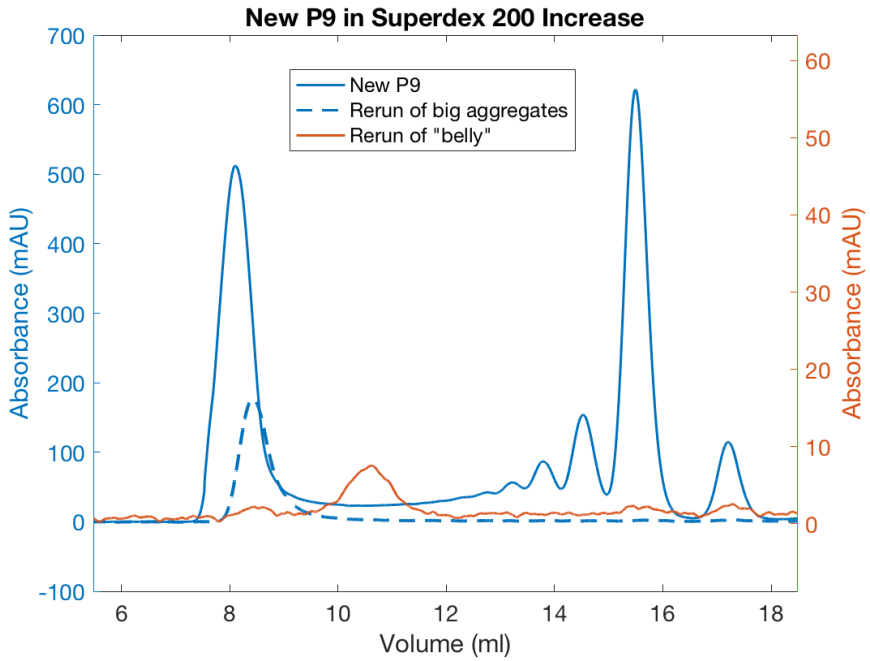

Figure 36: Absorbance (214 nm light) of eluate of the new P9, separated in Superdex 200 Increase. The proteins were incubated for one hour at protein concentration 1mM, in 2 mM HEPES buffer, pH 7.5, at 120% Ca^{2+} saturation and at room temperature. 3 hours later, fractions of the large aggregates (eluate volume of 7.5-8 ml) respectively of the belly (eluate volume of 10-10.5 ml) were analysed again with the same column. The re-run of the belly fraction is plotted to an other scale, following the right y-axis (red).

One last try to separate the particles in the void volume was made in a Superose 6 Increase volume. The result from this is shown in Figure 37. Still, many of the particles are too large to be separated in this column and are hence eluted in the void volume. Furthermore, a broad and low hump are seen in the belly, and the smaller aggregates are united into one single peak, since this column did not have high enough resolution to separate these.

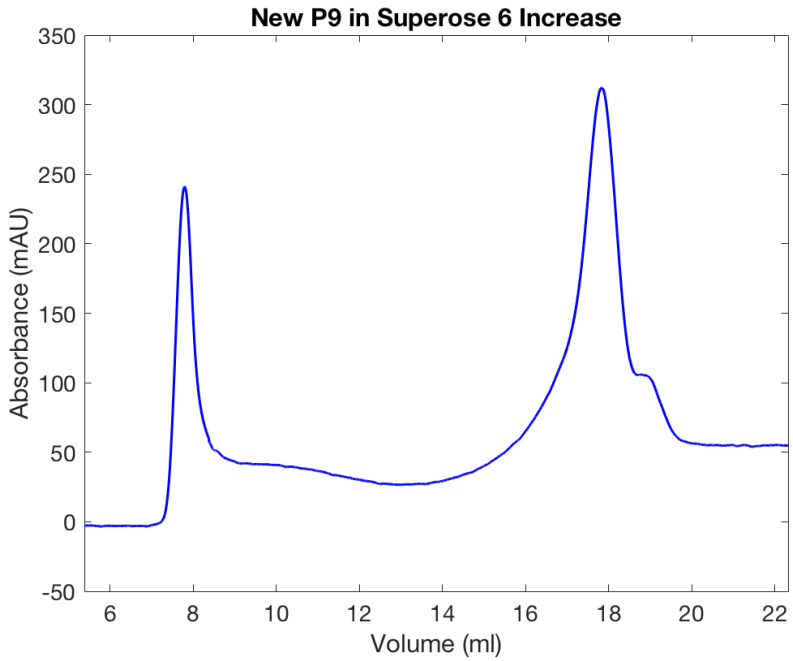

Figure 37: Absorbance (214 nm light) of eluate of the new P9, separated in Superose 6 Increase. The proteins were incubated for one hour at protein concentration 1mM, in 2 mM HEPES buffer, pH 7.5, at 120% Ca^{2+} saturation and at room temperature.

To analyse if the large particles in the void volume were proteins and not DNA, for instance, an absorbance spectrum was measured of a fraction of this peak, after separation with Superose 6 Increase. Also, this was done with a fraction of the peak corresponding to the smaller aggregates (eluate volume of 18 ml). These absorbance spectra are shown in Figure 38. From this, it can be seen that the absorbance is very similar between the two compounds, indicating that they both consists of proteins.

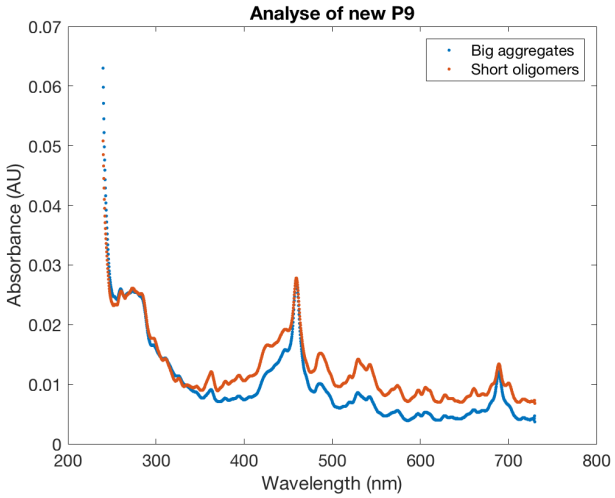

Figure 38: Absorbance spectra of the eluate containing the large particles from the void volume in Superose 6 Increase respectively of the short oligomers at the eluate volume of 18 ml in Figure 37.

4 Discussion

Each section will firstly present some reasoning regarding the experiments and the results, and end with the answer to the question(s) which the assay had as main aim(s) to answer.

4.1 Determination of Substrate Concentrations and Ca^{2+} Affinity Constants

One way to determine the concentrations of the Ca^{2+} additions, the Quin-2 and the proteins, relies on an absorbance measurement of the Quin-2. Another way is with the weight of $CaCl_2$ and its dissolution with water. Optimally, these concentrations would be equal, but this is not the case in this assay. The absorbance measurement gave a Ca^{2+} concentration of 2.5 mM and the weight of $CaCl_2$ gave 3.0 mM Ca^{2+}. But which method gives the concentration closest to the truth? For the latter method, one major source of errors is that $CaCl_2$ is hygroscopic and thus the molecular weight will differ depending on how much water the substance has absorbed when it is weighed. $CaCl_2$ needs to be stored in tightly sealed containers to keep the water level from rising. For the Ca^{2+} concentration calculations, each $CaCl_2$ was assumed to bind two water molecules (the chemical where produced to initially have this water level), giving the molecular weight of 147.02 g/mol. But if the molecules actually bound four water molecules, the molecular weight is 183.05 g/mol (Ropp, 2013), meaning a 80% lower Ca^{2+} concentration. 2.5 mM is 83% of 3.0 mM, so this could be an explanation.

On the other hand, the extinction coefficient used to calculate the Quin-2 concentration affects the calculated concentration linear, becoming a critical parameter. The extinction coefficient at the wavelength 239.5 nm is 42 000 cm^{-1} M^{-1} according to Linse et al., 1987. This value was obtained using absorbance measurements with known concentration of Quin-2, but it is hard to tell how accurate the concentrations were known in that study.

Based on above reasoning, it is likely that the Ca^{2+} concentration lies somewhere between 2.5 and 3.0 mM. For the experiments that followed, mean values of the obtained protein concentrations were used. In this project, the protein concentration is not needed to be very exactly determined for the overall goal: to find the best conditions of aggregation. Hence, it is sufficient to know the protein concentrations approximately.

The methodology of the Ca^{2+} titrations follows earlier experiments by Linse, 2002, except that instead of manual titrations, now the Probe Drum instrument was used to automate the process. Linse, 2002, states that a suitable time between addition of $CaCl_2$ and absorbance measurement (the equilibrium time) is 20-60 seconds. In this project, 30 seconds were used for the G-series, but due to some uncertainty regarding the slope at saturation, it was decided to increase this to 45 seconds for the P-series. The necessary time for the Ca^{2+}, chelator and protein to be in equilibrium, could be analysed with an assay where absorbance measurements are done frequently after an addition of $CaCl_2$, until no change of absorbance can be detected. The time this takes can be used as equilibrium time for a titration assay.

From Figure 8 and Table 2, it is clear that S100G has similar affinity for Ca^{2+} when inserting different lengths of polyglycine linkers. For the G-series, the average minimum cooperativity in this project is -9.6 kJ/mol and the average $\log K_1$ and $\log K_2$ are 7.7 respectively 8.7. These values can be compared to a previous study where the average minimum cooperativity for S100G with polyglycine linkers of 1-16 in length was shown to be -8.6 ±

0.1 kJ/mol and the average binding constants to be 7.79 ± 0.08 for $\log K_1$ and 8.66 ± 0.06 for $\log K_2$ (Linse, et al., 2020). Since the experiment of this project only considers three of the linker lengths, the mean values should not be compared straight away, but can be used as an indication for reliability assessment. Based on this, the assay of this project is considered to give reliable results.

Do the Ca^{2+} affinity of S100G significantly change depending on if the inserted linker consist of prolines or glycines? No, the total affinity is kept constantly high, regardless of the inserted linker. This is important since high Ca^{2+} affinity is needed for the Ca^{2+} responsiveness of a later designed material. However, the cooperativity decreases with inserted polyproline linker, especially so for 7 and 9 proline residues long linkers.

4.2 Dynamic Light Scattering

The serial dilution with Ca^{2+}-free buffer to the saturation series contributes with some additional Ca^{2+}, but the error is small enough to be ignored at this screening stage of analysis. The error is largest for 2.5 µM at 50% saturation, where the saturation instead becomes (2.5+0.3)/5 = 56%).

For P5 in the Ca^{2+}-free buffer and at 50% Ca^{2+} saturation, a clear trend can be seen; the decay slope is located at longer delay times the longer the time after protein dissolutions. From a qualitative point of view, this could be seen as an indication of aggregation into larger particles as the time goes. But, it is possible to calculate particle sizes from fittings to exponential functions for each data series, and the sizes obtained with this method tells that the particles corresponding to the samples after two, five and six days are much larger than should be possible since the samples were filtered prior to the measurement of the second day.

The fittings were made in MATLAB using three exponential terms, on the form: $f(x) = a \cdot e^{-bx} + c \cdot e^{-dx} + e \cdot e^{-fx} + g$ and the size of particles that each term correlates to was calculated with the exponential coefficient, b, d or f above and called b in below formula for the hydrodynamic radius, R_H (where Stokes-Einstein´s equation is used):

$$R_H = \frac{2kT}{(6\pi\eta b)q^2} \tag{4}$$

k is Boltzmann´s constant, T is the temperature in Kelvin, η is the viscosity of the medium and q is the scattering vector $q = 4\pi n/\lambda sin(\theta/2)$, where n is the medium refraction index, λ is the wavelength of the light and $\theta = 90°$ is the scattering angle.

The used filters had pores that should let particles smaller than 100 kDa pass, which corresponds to a minimum radius of 3.05 nm (Erickson, 2009). Thus, after the spin-filtration, no particles larger than about 3 nm in radius should be present in the samples. But the fittings give particle sizes of everything from 1.3 nm to 60 nm, indicating that something is wrong here.

Also the software for the instrument generated fittings to the data and calculated particle sizes, but no consistency of this values could be found, even when the Regularization fitting was used, which is the one to use for multimodal samples.

For the samples with 90% and 120% Ca^{2+} saturation, the data of the two hours measurement looks like it corresponds to larger particles than after 1 day, and smaller than after

two days, but the same reasoning applies for this results as for the previously analysed saturation levels; the autocorrelation slopes of measurements after two, five and six days corresponds to much too large particles to be possible after filtration.

The conclusion is that the DLS seems to detect some particles but does not have the resolution to compare sizes of aggregates. At least not at the edge of its detection limit of size, which is 1 nm in radius, and also for multimodal samples (Filipe et al., 2010).

Considering the inconsistent results, the fact that S100G has a size that is on the edge of detection limit for the DLS instrument, and that the theory suggests that DLS is not very accurate in multimodal solutions, this method is considered to be an unsuited choice for analysing the aggregation. Hence, no more focus will be given the DLS in this report.

Is DLS a suitable method for analysing oligomeric states of S100G, and if so, in what calcium and protein concentrations do the proteins aggregate the most? No, DLS seems not to be suitable for distinguish between oligomers on this size scale.

4.3 Native Gel Electrophoresis

Since no aggregates were detected in Ca^{2+} saturation levels other than over 100%, it is assumed that the formed aggregates contains Ca^{2+} with the molar ratio of saturation, 1:2.

Since only P5 was used in the screening of concentrations and Ca^{2+} saturation levels, one might wonder if the conclusions of this assay directly can be translated to the other substrates? With most certainty this will be the case for the protein concentration, since the substrates are relatively similar. But the Ca^{2+} concentration might be more dependent of the linkers, and from Linse et al., 2020, it was found that the G-series created significant oligomeric structures at 80% Ca^{2+} saturation. This is the reason both 80% and 120% saturation were used in the later PAGEs.

However, the G-series did not show any tendencies to form higher oligomeric structures than dimers, in both Ca^{2+} saturation levels. This is in contrast with the aggregation that was found in Linse et al., 2020, which showed up to seven variants of oligomers for G12. But, it shall be noted, that assay was performed at pH 5 and 6, with 0.1 M NaCl. These could be explanations for the difference in the results.

Since the G-series was not under optimal conditions to aggregate in this assay, it can not be ruled out that a combination of proteins from the G-series and P-series could aggregate successfully together at other conditions. The P-series showed aggregation in a broader range of conditions (showed in the SEC assays) and hence, it might have been another result if the conditions were changed to favour the G-series more. Håkansson, et al., 2001, suggests that a pH around 5 is necessary for S100G P43M to undergo domain swapping, and since Linse et al., 2020, tells this variant of S100G behaves similar to the G-series, it is likely that the pH is a reason for the absent oligomerisation of the G-series in this project.

It was believed but not evaluated, that a higher Ca^{2+} saturation level than 120%, would not be more, or less, beneficial for aggregation. It was thought to be unlikely that the aggregation would behave differently in this concentration region since all proteins were saturated with marginal already at 120%. This is the reason for not performing the assay evaluating this earlier in the project. But since there was place over in the gel used to see how the newly desalted P9 behaved, an assay which could support this theory was easy to set up. The result supported the theory; more than 120% did not have significance impact of the aggregation.

Worth mentioning when analysing this last gel, is that the loading dye was seen to travel downwards during the run in a slightly different way. It is supposed to travel with the same speed no matter the location, but in this gel, the dye in the wells closest to the edges was a bit slower, which tells that when comparing the bands height in the result, one should keep in mind that the very same protein in all wells would give rise to a slightly bend line; quite straight in the middle and more and more bent upwards the closer the edges the well lies. Especially in comparison of well 1-4 (the new P9 and the old in different Ca^{2+} saturation levels) this is important.

Is it possible to detect any oligomeric states in a native PAGE, and if so, at what protein and calcium concentrations are this happening? Yes it is possible. The best conditions to detect oligomers in a native PAGE are with a protein concentration of about 1 mM and at a Ca^{2+} concentration well above saturation.

At 80% respectively 120% Ca^{2+} saturation, how does the linker length and type affect the aggregation, and is it more beneficial to combine different substrates? An inserted linker with polyprolines gives more oligomeric structures than with glycines. P9 had the most different states, with up to three oligomeric structures detected, in addition to the monomer. No additional oligomeric states arises when mixing different substrates, at least not at pH 7.5.

Is the aggregation dependent on Ca2+concentration above 120% saturation? No, no such dependence above 120% saturation could be found.

Have the aggregates changed after 50 days of incubation? For G5, G7 and G12, no significant change could be seen except for some protein degradation to smaller parts. P9 seems to have lost some larger oligomeric structures.

4.4 MDS

From diffusion measurements, the mean hydrodynamic radius can be obtained, which is not possible in a Native PAGE. Diffusion is a good complement to native PAGE because it does not consider the charge in the measurements, making it possible to strengthen the indications that the proteins with polyproline linkers with 120% Ca^{2+} saturation gives more aggregation than at lower saturation levels.

Dilutions prior to addition to chip were made by adding 1 µL of 1 mM protein solution, to 24 µL HEPES Ca^{2+}-free buffer. It is possible that pipetting 1 µL might give a significant error in volume added, and this might contribute to the difference in calculated concentrations. Also, the amount of protein in each aliquot might have some error, since this comes from the mean concentrations calculated using Table 2. However, since the concentrations are in the region that the instrument can work in, the difference is not important for the hydrodynamic radius.

No aggregates could be found in the native PAGE assay for substrates in Ca^{2+}-free buffer, telling that it is likely that the analysis with MDS of the substrates in Ca^{2+}-free buffer measures monomeric proteins. The mean radii after one day of incubation was 1.17 nm for P5, 1.57 nm for P7 and 1.34 nm for P9. What is the reason for the large gap in size between P5 and P7, and has P7 actually larger hydrodynamic radius than P9? In the topology analysis in the introduction section, S100G was found to be be approximately 25 Å in diameter and the hydrodynamic radius should be smaller than this. The linker length of Pro_5 is about 13.6 Å and Pro_9 about 24.3 Å. How much the inserted linkers affect the

hydrodynamic radii is hard to tell, it will depend on how the linker folds. But a P5 radius of 1.17 nm seems realistic or somewhat too small, and 1.57 nm for P7 might be too large. Since also SEC depend of the hydrodynamic radius, the results from Figure 26 can be compared with this estimated hydrodynamic radii. From this, it looks like P7 and P9 are a bit more similar in size than P5 and P7, but still P7 is smaller than P9. The conclusion is that the measured sizes with MDS should not be seen as absolute truths, but rather as indications of sizes.

The MDS size estimations of the monomeric proteins can be compared to the size measurements of the aggregates that were made using MDS in eluate from SEC. From Table 5 and 6, the hydrodynamic radius of each aggregate size is given. But in this analysis, only Table 6 will be used since many of the results in Table 5 could not be obtained due to instrument problems. The smallest particles corresponds to the last peak for each substrate. It can be seen that all of them are larger than their monomeric radii; third peak of P5 has 1.74 $\approx$ 1.7 nm in radius, fourth peak of P7 has 1.82 $\approx$ 1.8 nm and fourth peak of P9 has 1.88 $\approx$ 1.9 nm. Even considering some error in the measurements, this seems consistently larger than monomers. A lower estimation of how much larger a formed aggregate is compared to a monomer (radius r) can be obtained with the approximation that the proteins fuse into one perfect sphere with the radius R. From the volume ratios of the spheres, the radius ratio can be derived to:

$$R/r = a^{-1/3} \tag{5}$$

where a is the number of proteins that build up the aggregate. A higher limit estimation of the ratio is obtained with an approximation that the proteins are perfect spheres and stacks in a linear way, without fusing at all, called the liner bead model in Adamczyk, et al., 2010. Ratios for up to ten proteins in one aggregates are given in Table 7, for both models.

The P5 (R/r) ratio of the third peak (smallest particles) is estimated to 1.7 nm/1.2 nm $\approx$ 1.4. For P7 it is 1.8 nm/1.4 nm $\approx$ 1.3 and for P9 it is 1.9 nm/1.4 nm $\approx$ 1.4. Based on this, it looks like the smallest particles in 120% Ca^{2+} saturation are dimers. But despite the fact that the ratios are correlating well, it is important to consider Figure 26, which shows eluate absorbance for all three substrates in the same plots as S100G. Since the literature (see the introduction section for more information) tells that S100G is not likely to form dimers at Ca^{2+} saturation, the high peak in this graph can be said to be monomers for sure. P5, P7 and P9 are shifted to the left of this peak, but not by much. Hence, the theory that this would be dimers can be ruled out. It remains to conclude that the measurements of the hydrodynamic radius performed with MDS has a large standard deviation and can not be trusted for further analysis in this project.

Table 7: Ratio (R/r) of aggregate radius and monomer radius, obtained with Equation 5 for a fused sphere model, respectively the linear bead model of Adamczyk, et al., 2010.

Number of proteins	2	3	4	5	6	7	8	9	10
Fused spheres	1.39	1.73	2.04	2.32	2.60	2.86	3.12	3.37	3.61
Linear bead model	1.26	1.44	1.59	1.71	1.82	1.91	2.00	2.08	2.15

Can MDS be used to support the native PAGE results? At 120% Ca^{2+} saturation, the hydrodynamic radii of P5 and P9 are larger than in calcium free solution, which is in line with the native PAGE results. But MDS seems to have a relatively large error in the size determinations, making the results unreliable.

What sizes have different oligomers of P5, P7 respectively P9? MDS seems to have a too large error in the size determinations to answer this question with certainty.

4.5 Assays that use SEC

The initial mistake to not have 2.4 mM $CaCl_2$ in the running buffer turned out to be valuable. As seen in the results of this assay, only two peaks of each substrate were found in the eluate, compared to three peaks for P5 and four for P7 and P9, which were the cases with 2.4 mM in the running buffer. Also, the belly between the two peaks indicate that some aggregates changed between the two states during traveling in the column, giving the mean velocity corresponding to somewhere in between the peaks. This instability of the aggregates was further confirmed when reruns of fractions of both peaks showed that all aggregates had been broken down to the smallest particles in the time scale of 30 minutes.

In contrast, with 2.4 mM $CaCl_2$ in the sample *and* running buffer, the aggregates were much more stable (Figure 20). In this case, it was P5 that was analysed with reruns, but since Figure 22 shows that P5 has a belly that is more shifted towards the second peak than for P9, it is safe to say that the same instability phenomena is happening for P5, if so only stronger. Two conclusions are drawn from this. Firstly, Ca^{2+} exchange is a fast process, meaning that the proteins will adopt to a new Ca^{2+} concentration. Secondly, the aggregates that are formed under Ca^{2+} saturated conditions will break down when the Ca^{2+} are released from the proteins. The Ca^{2+} adopting to a lower concentration might occur very fast and the dissociation of aggregates in a slower fashion, but from these experiments it can only be said that the total time of break down is on the timescale of tens of minutes, since it takes 10-13 minutes for the proteins to travel along the column.

The running buffer had the same content (except for the 0.1 M NaCl) as the proteins were dissolved in, and the column was filled with it prior to injections. Hence, the proteins were kept under similar conditions through the filtration process. Only dilution, ionic strength and perhaps interference with the stationary phase in the column should be parameters that differ. The Ca^{2+} concentration was 2.4 mM in both the sample and the running buffer, so when the proteins were diluted in the column, the Ca^{2+} saturation level increased, but according to the native PAGE where more than 120% Ca^{2+} saturation was evaluated, this does not matter significantly.

For different pH of the samples, the same running buffer was used (pH 7.5). This was done since it was not possible to have a running buffer equal to the protein solution buffer, since the SEC do not work good under all conditions, and also it would have taken very long time to run the experiments if the buffer should be changed between each run. Therefore, it was decided to keep one running buffer for all samples, rather than change it sometimes to a more similar but still run-able buffer, since the same running buffer will allow for better comparison between experiments and give fewer parameters to consider.

So, how much does this matter for the size exclusion chromatography? Firstly, it should be said that even if it is fast-forming aggregates, they seem to be stable in dilution at lower protein concentration, shown in experiments with re-runs of fractions. Hence, the dilution in the running buffer should not be a major problem on this timescale (10-15 minutes),

as long as the Ca^{2+} saturation level is kept constantly high. But of course it can not be said that the same stability applies for change of pH. This is the reason the assay with pH 5 of both column and protein sample was performed. The result of this assay shows that there seem to be differences using pH 5 or 7.5 in the column, when the sample has pH 5. However, from Figure 33, it looks like there should be almost no difference at all in aggregation between pH 5 and 7.5.

An explanation might be that the pH changes the Ca^{2+} affinity (Håkansson et al., 2001; Kesvatera et al., 2001), affecting the Ca^{2+} release and uptake, which is a fast process. Hence, it does not matter what pH the sample where in before injection, the Ca^{2+} affinity will fast adapt to the column pH and aggregate according to how much protein is in complex with Ca^{2+}. Thus, the pH evaluation assay might not at all be a proof of a low pH dependence, since this would require a running buffer of the same pH. But, one should not forget that it is not a radically big difference between the graphs in Figure 29, so at least between 5 and 7.5 in pH, the aggregation is not changing drastically.

Since also the ionic strength affects the Ca^{2+} affinity (Linse et al., 1991), one might argue that the logic of the fast adopting aggregation states should apply to ionic strength as well. But there is a significant difference between the eluate spectra with and without 0.15 M NaCl, which is not the case with varying pH and where the running buffer had a constant pH, as discussed above. This indicate that at different ionic strengths, something more that Ca^{2+} affinity changes. It fits well with theory stating that the hydrophobic force and interactions between EF-hands are stronger at 0.15 M NaCl than without. This might be a reason the aggregates are more stable with changing salt concentration than with changing pH.

In the assay evaluating the pH and ionic strength dependence, the temperature was chosen to be 50°C. There is a probability that other combinations of temperatures, pH and ionic strength could give more aggregation, but since it was not possible to try all the different combinations, the most promising conditions in previously assays were decided to be used for the coming ones. The same logic was used when the Ca^{2+} saturation level was evaluated; the best level was used in further studies. Also, other variants of S100G could behave differently. The optimal conditions for aggregation in this project should therefore be considered as a good point to start a further work, aiming to form a material, but not as the absolute best condition.

There were quite significant differences in the signal strengths (around a factor 2) of the SEC results that were used to generate Figure 23-25, which can be seen in the non-normalized plots in Appendix D, Figure 49-51. The reason for this is that the amount of substrate analysed with SEC, varied somewhat. 50 µL were always aimed at, but due to the small volumes of the samples, it was hard to dispose the sample volume correctly among the measurements. Also, some of the measurements were performed with another brand of FPLC machine and this had an other type of syringe, which made the loading volume a bit less exact. However, in the qualitative analysis that is done in this report, the variance of loaded volume should not affect any conclusions, since the profile for each absorbance is the same independently of signal amplitude. Thus, this is not considered to be any significant source of error.

It can be noted that, under all the different conditions the P9 substrates were incubated in, and during various times, all spectra are relatively similar in shape. Yes, there are differences, which are pointed out and analysed in the result section, but considering the big difference of conditions (pH-range of 5-9, 0 M respectively 0.15 M NaCl, 4°C-70°C, and

incubation time of a few minutes to one week), it would not be surprising if some conditions had given no aggregation at all or much larger aggregates. Hence, it should be said that the distributions of the aggregates are overall quite constant and not very dependent of the conditions in the given ranges.

Are the aggregates fast forming and reforming, and how stable are they diluted? The aggregates are very fast forming, on a timescale of a minute. They are not reforming into other sizes significantly much on a timescale of at least 30 minutes, even diluted. But if the Ca^{2+} concentration decreases below saturation, the aggregates dissociate on a timescale of ten minutes. The dimeric structure seems to be the most stable non-monomeric form.

How do the aggregates change with the time since dissolutions? In the time interval of a few minutes up to one week, the oligomers differs only slightly. The dimeric form is most stable and for both P7 and P9, this form increases in numbers over time.

How does the temperature affects what aggregates are formed? For P9, between 4°C and 50°C, there is no or almost no temperature dependence of the aggregation. At 70°, larger aggregates are less common.

How does the pH affects what aggregates are formed? The pH affects the Ca^{2+} affinity and thus, the aggregation. This is a fast response, so the running buffer in a SEC needs to have a similar pH as the sample to detect any differences in aggregation. At pH 5, less aggregation occur for P9, than at 7.5, but still all three oligomeric states are present, in addition to the monomers.

How does the addition of 0.15 M NaCl affect what aggregates are formed? The same aggregates of P9 are still formed in this ionic strength, but the larger ones seems to break down to dimers faster than without added salt.

4.6 Comparison of Newly Desalted P9 with Earlier Used P9 Aliquots

From the native PAGE, where this issue was regarded, it looked like the new P9 formed similar oligomers as the old P9, only one or two more different states. But the analysis with SEC tells a different story. Here, it can be seen that the new P9 aggregates into much larger particles. So large that even the Superose 6 Increase column could not separate them, which is the column able to separate the largest particles of the used ones. Its exclusion limit (that is, the particle size which does not penetrate the resin pores, making them being eluted in the void volume) corresponds to about 40 MDa globular proteins. With a density of 1.37 g/cm^3, this is about 23 nm in radius. Given an estimated monomeric radius of 1.5 nm, this corresponds to particles with radii about 15 times the monomeric one. It may also be noted that the particles were small enough to pass the column filter. From the comparison of absorbance spectra in Figure 38, it can be ruled out that the larger particles are DNA from the purification, since the absorbance of both spectra are very similar in the interval 250-300 nm.

The shift of about 0.5 ml for the re-run peak of the largest particles in Figure 36, is likely due to the larger loop used for the re-runs. Only 100 µL of the total loop volume of 500 µL were used here. This is in contrast with the re-run of the fraction of the belly, where the 500 µL loop was filled to get a higher signal, and thus this peak is not shifted. This tells that the belly is not a result of particles that change their sizes along the column. Instead, it arises because there are proteins with sizes corresponding to these eluate volumes.

From this analysis, it can be concluded that the new P9 forms particles that are larger than the corresponding size of a 40 MDa globular protein and they do not break down significantly in 3 hours time. Also, particles from this size down to short oligomers, are formed and stable in at least 3 hours time. To get knowledge about how discrete these aggregate sizes are, more experiments are needed, with higher resolution in the separation of this sizes.

It may be noted that when using the column Superose 6 Increase, the absorbance of 214 nm does not reach the zero base-line after the last peak. Since both the base-line before the first peak and after the last one are flat, it is not likely air bubbles in the detector, but some other artifact in the measurement. Since the other peaks can be described with earlier results, they seem to be reliable even though the baseline has increased toward the end of the graph.

Do the newly desalted P9 aggregate in a similar way as the earlier used P9? In a native PAGE, the new P9 forms similar oligomers, but it has one or two higher oligomeric structures. In SEC, the new P9 forms much larger aggregates.

5 Conclusions

The results of this study show that S100G, with inserted polyproline linkers, is able to form oligomers of higher order than dimers, in presence of a Ca^{2+} concentration that well saturates the proteins. With nine proline residues inserted in the linker loop (P9), the highest number of different oligomeric structures are formed, with up to three different structures in addition to the monomers. Another fraction of P9 from the purification process, formed much larger aggregates than this, with sizes from dimers up to particles with radius more than 15 times the monomeric radius. But more analysis of this aggregates is needed to understand its structure and why this P9 behaves differently from the other.

The optimal conditions for oligomerisation of the first analysed batch of P9, are under well saturated Ca^{2+} concentrations (120% was used), pH 7.5 and at 50°C (careful not to go higher). But all oligomeric states are still present to some extent at pH down to 5, temperatures down to 4°C, respectively with addition of 0.15 M NaCl.

If the conditions are changed to lower Ca^{2+} concentrations than saturation, the aggregates will break down significantly in 10 minutes and totally to monomers within one hour. Hence, changing a condition that directly affects the Ca^{2+} affinity, such as pH, will affect the aggregation. The dimeric forms seems to be the most stable aggregate, considering time, pH, ionic strength and temperature.

The Ca^{2+} affinity is kept at an equal or almost equal high level when insertion of 5-9 proline residues or 5-12 glycine residues between the EF-hands. But the cooperativity decreases drastically if it is polyprolines compared to polyglycines.

6 Further Development of the Project

Below are some areas listed that could be of interest to look further into as a continuation to this project, with the final goal of designing a Ca^{2+} responsive reticular material.

- Further analysis of the other P9 batch needs, to understand what differs between this protein and the P9 used earlier in this project. Some amino acid might differ and this may give hints for designing further candidates of a Ca^{2+} responsive reticular material. For instance, circular dichroism could be used to see if the secondary structure is varying among the aggregates.
- Substrates of S100G EF-hand 1, linked by polyproline in various lengths to another EF-hand 1, have been synthesised prior to this project. This is also the case for EF-hand 2 connected to another EF-hand 2. Due to lack of time, this substrates were not analysed, but it could be of interest to mix these substrates with each other, and with the substrates used in this project, to see if more aggregation can be provoked in this way.
- Based on the results of this project, other variants of the substrate S100G can be used under similar conditions. For instance, it would be of interest to produce substrates with three EF-hands and with linkers of polyproline. Also, other proteins in the EF-family, can be used to see if they are more suitable for aggregation into more ordered structures. To create a specific form of a reticulum, the linkers can be designed to have certain angles that are preferred, for instance by introducing a kink in a helical structure.
- Once a material is formed, suitable molecules can be added and see if they can be encapsulated in the structure, generating a host-guest molecular structure.
- To solve the oligomeric structures at atomic resolution, X-ray crystallography can be used. This might provide important hints on how to continue the designing of a Ca^{2+} responsive reticular material. Knowledge, such as what molecular parts are prone to aggregation and what functional groups are involved in both the binding and at the surface, can be of value for predictions of suitable conditions and potential guest molecules in a future material. Also methods such as NMR, ultracentrifugation and CD spectroscopy can be used to evaluate what structures are formed upon aggregation.
- An assay can be performed where substrates from G-series and P-series are mixed, under the same conditions as Linse et al., 2020 (supplementary information) had for the G-series.

References

Adamczyk, Z., Sadlej, K., Wajnryb, E., Ekiel-Jezewska, M. L., Warszynski, P. 2010. *Hydrodynamic Radii and Diffusion Coefficients of Particle Aggregates Derived from the Bead Model*. Elsevier. Journal of Colloid and Interface Science 347: 192–201. DOI: 10.1016/j.jcis.2010.03.066.

Akke, M., Forsén, S. (1990). Protein Stability and Electrostatic Interactions Between Solvent Exposed Charged Side Chains. *PROTEINS: Structure, Function, and Genetics*, 8:23-29. DOI: 10.1002/prot.340080106.

Akke, M., Drakenberg, T., and Walter J.C. (1992). Three-Dimensional Solution Structure of Ca2+-Loaded Porcine Calbindin D9k Determined by Nuclear Magnetic Resonance Spectroscopy. *Biochemistry*, 31:1011-1020. DOI: 10.1016/0022-2836(91)90389-n.

Andersson, M., Malmendal, A., Linse, S., Ivarsson, I., Forsen, S., and Svensson, L.A. (1997). Structural Basis for the Negative Allostery Between Ca2+- and Mg2+-binding in the Intracellular Ca2+-Receptor Calbindin D9k. *Protein Science*, 6:1139-1147.

André, I., Linse, S. (2002) Measurement of Ca^{2+}-Binding Constants of Proteins and Presentation of the CaLigator Software. *Analytical Biochemistry*, 305:195–205. DOI: 10.1006/abio.2002.5661

Berggård, T., Julenius, K., Ogard, A., Drakenberg T., and Linse, S. (2001). Fragment Complementation Studies of Protein Stabilization by Hydrophobic Core Residues. *Biochemistry*, 40:1257-1264.

Bootz, A., Vogel, V., Schubertb, D., and Kreuter, J. (2004). Comparison of Scanning Electron Microscopy, Dynamic Light Scattering and Analytical Ultracentrifugation for the Sizing of Poly(butyl cyanoacrylate) Nanoparticles. *European Journal of Pharmaceutics and Biopharmaceutics*, 57(2):369–375.

Choi, K. C., and Jeung, E. B. (2008). Molecular Mechanism of Regulation of the Ca^{2+}-Binding Protein Calbindin-D9k,and its Physiological Role(s) in mammals: a Review of Current Research. *Journal of cellular and molecular medicine*, 12:409–420.

Clark, A.H., and Ross-Murphy, S.B. (2009). *CHAPTER 1 - Biopolymer Network Assembly: Measurement and Theory*, In: Kasapis, S., Norton, I.T., Ubbink, J.B. (Editors) *Modern Biopolymer Science*. Academic Press. ISBN 9780123741950.

Dell'Orco, D., Xue, W.F., Thulin, E., and Linse, S. (2005). Electrostatic Contributions to the Kinetics and Thermodynamics of Protein Assembly. *Biophysical Journal*, 88:1991-2002.

Erickson, H.P. (2009). *Size and Shape of Protein Molecules at the Nanometer Level Determined by Sedimentation, Gel Filtration, and Electron Microscopy*. Biological Procedures Online, Volume 11, Number 1. DOI: 10.1007/s12575-009-9008-x

Finn, B.E., Kördel, J., Thulin, E., Sellers P., and Forsén, S. (1992). Dissection of Calbindin D9k Into Two Ca2+ Binding Subdomains by a Combination of Mutage-

nesis and Chemical Cleavage. *Federation of European Biochemical Societies*, 298:211-214.

Filipe, V., Hawe, A., and Jiskoot, W. (2010). Critical Evaluation of Nanoparticle Tracking Analysis (NTA) by NanoSight for the Measurement of Nanoparticles and Protein Aggregates. *Pharmaceutical research*, 27(5):796–810. DOI: 10.1007/s11095-010-0073-2

Fluidic Analytics. *Microfluidic Diffusional Sizing for protein analysis; an introduction*. Viewed 03-05-2021, available at: `https://bitesizebio.com/wp-content/uploads/2019/03/Bitesize-Bio-Fluidic-Analytics-download.pdf`

Guo, Z., and Eisenberg, D. (2006). Runaway Domain Swapping in Amyloid-Like Fibrils of T7 Endonuclease I. *PNAS*, 21(103):8042–8047. DOI: 10.1073.pnas.0602607103

Håkansson, M., Svensson, A., Fast, J., and Linse, S. (2001). An extended hydrophobic core induces EF-hand swapping. *Protein Science*, 10:927-933.

Jaskolski, M. (2001). 3D Domain Swapping, Protein Oligomerization, and Amyloid Formation, *Acta Biochim Pol.* 48:807–827.

Kesvatera, T., Jönsson, B., Telling, A., Tõugu, V., Vija, H., Thulin, E., and Linse, S. (2001). Calbindin D9k: A Protein Optimized for Calcium Binding at Neutral pH. *Biochemistry*, 40:15334-15340.

Kördel, J., Skelton, N.J., Akke, M., Chazin, W.J. (1993). High-resolution Solution Structure of Calcium-loaded Calbindin D9k *Journal of Molecular Biology*, 231:711-734

Lawson, C.L., Benoff, B., Berger, T., Berman, H.M., and Carey, J. (1993). E. coli trp Repressor Forms a Domain-Swapped Array in Aqueous Alcohol. *Protein Science*, 12:1099–1108.

Leukert, N., Vogl, T., Strupat, K., Reichelt, R., Sorg, C., Roth, J. (2006). Calcium-dependent Tetramer Formation of S100A8 and S100A9 is Essential for Biological Activity, *Journal of Molecular Biology*, 359(4):961-972.

Li. C., Arakawa, T. 2018. *Application of Native Polyacrylamide Gel Electrophoresis For Protein Analysis: Bovine Serum Albumin as a Model Protein.* International Journal of Biological Macromolecules 125:566–571. DOI: 10.1016/j.ijbiomac.2018.12.090

Linse, S., Brodin, P., Drakenberg, T., Thulin, E., Sellers, P., Elmdén, K., Grundstrom, T., and Forsen, S. (1987). Structure-Function Relationships in EF-Hand Ca2+-Binding Proteins. Protein Engineering and Biophysical Studies of Calbindin D9k.*Biochemistry*, 1987, 26:6723-6735.

Linse, S., Teleman, O., and Drakenberg, T. (1990). Ca2+ Binding to Calbindin D9k Strongly Affects Backbone Dynamics: Measurements of Exchange Rates of Individual Amide Protons Using ^{1}H NMR. *Biochemistry*, 29:5925-5934.

Linse, S., Johansson, C., Brodin, P., Grundström, T., Drakenberg, T., and Forsén, S. (1991). Electrostatic Contributions to the Binding of Ca2+ in Calbindin D9J. *Biochemistry*, 30:154-162.

Linse, S., Thulin, E., and Sellers, P. (1993). Disulfide Bonds in Homo- and Heterodimers of EF-hand Subdomains of Calbindin Dgk: Stability, Calcium Binding, and NMR Studies *Protein Science*, 2:985-1000.

Linse, S. (2002). *Chapter 2 - Calcium Binding to Proteins Studied via Competition with Chromophoric Chelators.* In: H. J. Vogel (editor) *Methods in Molecular Biology", vol. 173: Calcium-Binding Protein Protocols, Vol. 2: Methods and Techniques*, Humana Press Inc.

Linse, S., Thulin, E., Nilsson, H., and Stigler, J. (2020). Benefits and Constrains of Covalency: the Role of loop Length in Protein Stability and Ligand Binding. *Scientific Reports*, 10:20108.

Madadlou, A., O'Sullivan, S., Sheehan, D. (2016). *Protein Chromatography - Methods and Protocols, Second Edition.* Humana Press, New York, NY. ISBN: 978-1-4939-6412-3.

Marenholz, I., Heizmann, C.W., and Fritz, G. (2004), S100 Proteins in Mouse and Man: From Evolution to Function and Pathology. *Biochemical and Biophysical Research Communications*, 322:1111–1122.

Malashkevich V.N., Dulyaninova N.G., Ramagopal U.A., Liriano M.A., Varney K.M., Knight D., Brenowitz M., Weber D.J., Almo S.C., and Bresnick A.R. (2010). Phenothiazines Inhibit S100A4 Function by Inducing Protein Oligomerization. *Proceedings of the National Academy of Sciences of the United States of America*, 107(19):8605-8610. DOI: 10.1073/pnas.0913660107.

Melnichenko, Yuri B. *Small-Angle Scattering from Confined and Interfacial Fluids.* (2016). Springer. DOI 10.1007/978-3-319-01104-2

Moroz, O.V., Antson, A.A., Dodson, E.J., Burrell, H.J., Grist, S.J., Lloyd, R.M., Maitland, N.J., Dodson, G.G., Wilson, K.S., Lukanidin E., and Bronstein, I.B. (2002). The structure of S100A12 in a hexameric form and its proposed role in receptor signalling. *Acta Crystallographica*, D58:407-413 DOI: 10.1107/S0907444901021278

Ostendorp, T., Leclerc, E., Galichet, A., Koch, M., Demling, N., Weigle, B., Heizmann, C. W., Kroneck, P. M., and Fritz, G. (2007). *Structural and functional insights into RAGE activation by multimeric S100B.* The EMBO journal, 26(16), 3868–3878. DOI: 10.1038/sj.emboj.7601805

Persechini, A., Moncrief,N.D. and Kretsinger, R.H.(1989). The EF-hand Family of Calcium-Modulated Proteins. *Elsevier Science Publishers*, 11(12):462-467.

Ropp, R.C. (2013). *Chapter 2 - Group 17 (H, F, Cl, Br, I) Alkaline Earth Compounds.* In: Ropp, R.C. *Encyclopedia of the Alkaline Earth Compounds*, Elsevier B.V. DOI: 10.1016/B978-0-444-59550-8.00002-8

Schlunegger, M.P., Bennett, M.J., and Eisenberg, D. (1997). Oligomer Formation By 3D Domain Swapping: A Model for Protein Assembly and Misassembly. *Advances in Protein Chemistry*, 50:61-122.

Seiffert, S. (2015). *Supramolecular Polymer Networks and Gels*, Springer. DOI: 10.1007/978-3-319-15404-6.

Stetefeld, J., McKenna, S.A., and Patel, T. R. (2016). Dynamic Light Scattering: A Practical Guide and Applications in Biomedical Sciences. *Biophysical Reviews*, 8:409–427.

Sun, C.Q., and Sun, Y. (2016). *Chapter 12 - Aqueous Solutions: Quantum Specification* In: *The Attribute of Water*. Springer Science+Business Media Singapore. DOI: 10.1007/978-981-10-0180-2_12

Swinehart, D. F. (1962) The Beer-Lambert Law, *Journal of Chemical Education* 39(7):333-335

ThermoFisher. *Native PAGE protein separation*. Viewed 13-05-2021, available at: `https://www.thermofisher.com/se/en/home/life-science/protein-biology/protein-gel-electrophoresis/protein-gels/specialized-protein-gels/nativepage-bis-tris-gels.html`

Thulin, E., Kesvatera, T., Linse, S. (2011). Molecular Determinants of S100B Oligomer Formation. *PLoS One*, 6(3):e14768. DOI: 10.1371/journal.pone.0014768 UniProt. *P29377 (S100G_HUMAN)*, Viewed 12-01-2021, available at: `https://www.uniprot.org/uniprot/P29377`

Wyatt Technology. (2014). *Dyna-LSU Course Manual*. Light Scattering University.

Wyatt Technology. *DynaPro Plate Reader - Specifications*. Viewed 14-05-2021, available at: `https://www.wyatt.com/products/instruments/dynapro-dynamic-light-scattering-plate-reader.html#dynaproplate-8`

Appendices

Appendix A - Ca^{2+} affinity

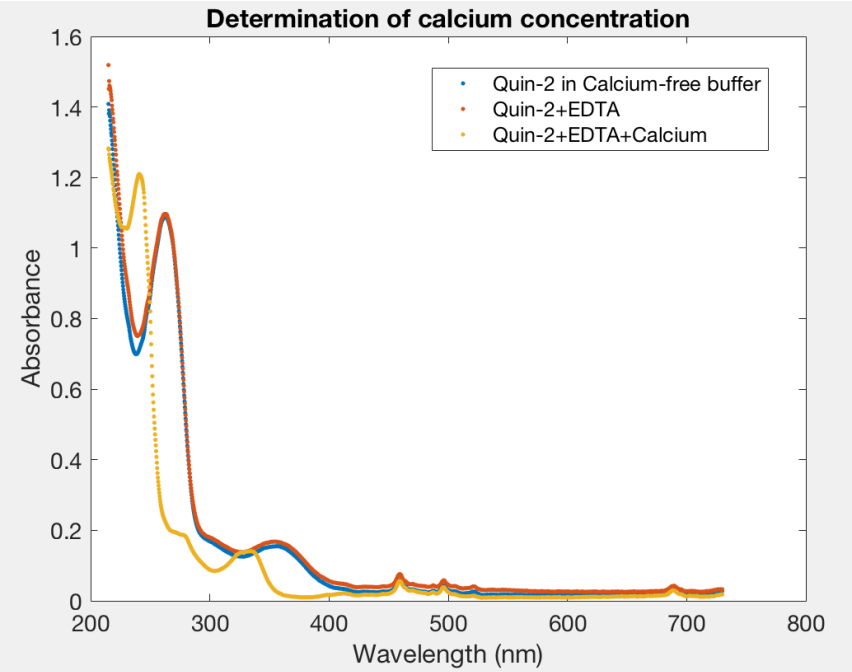

Figure 39: Absorbance spectra of 1 ml 28 µL Quin-2 in the Ca^{2+}-free buffer, with addition of first 2.5 µL 0.1 M EDTA, and finally also addition of 2.5 µL 1.0 M Ca^{2+}, which gives well saturated Quin-2 and EDTA. The absorbance of the wavelength 263 nm is changing strongly depending on Ca^{2+} saturation of Quin-2.

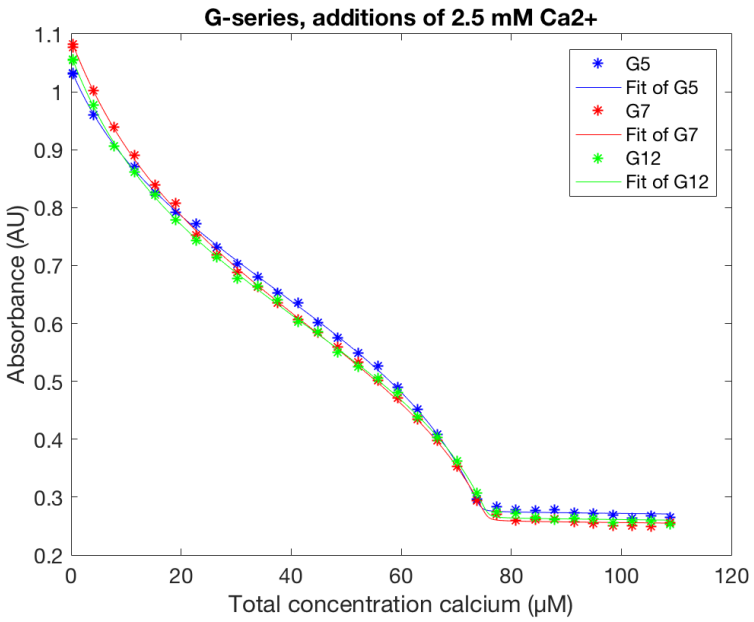

Figure 40: The absorbance of the wavelength 263 nm of the solution with both Quin-2 and the protein substrate G5, G7 or G12, plotted against the total concentration Ca^{2+} during the titration. The fitting from CaLigator is showed for each substrate. The concentrations used are 28.2 µL Quin-2 and 2.5 mM CaCl−2, in 2 mM TRIS buffer at pH 7.5.

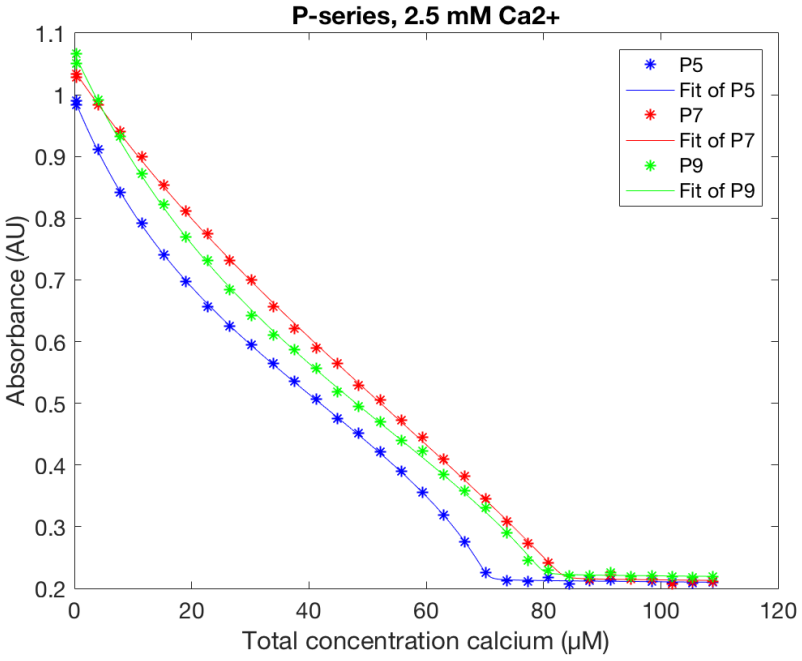

Figure 41: The absorbance of the wavelength 263 nm of the solution with both Quin-2 and the protein substrate P5, P7 or P9, plotted against the total concentration Ca^{2+} during the titration. The fitting from CaLigator is showed for each substrate. The concentrations used are 28.2 µL Quin-2 and 2.5 mM CaCl−2, in 2 mM TRIS buffer at pH 7.5.

Appendix B - DLS

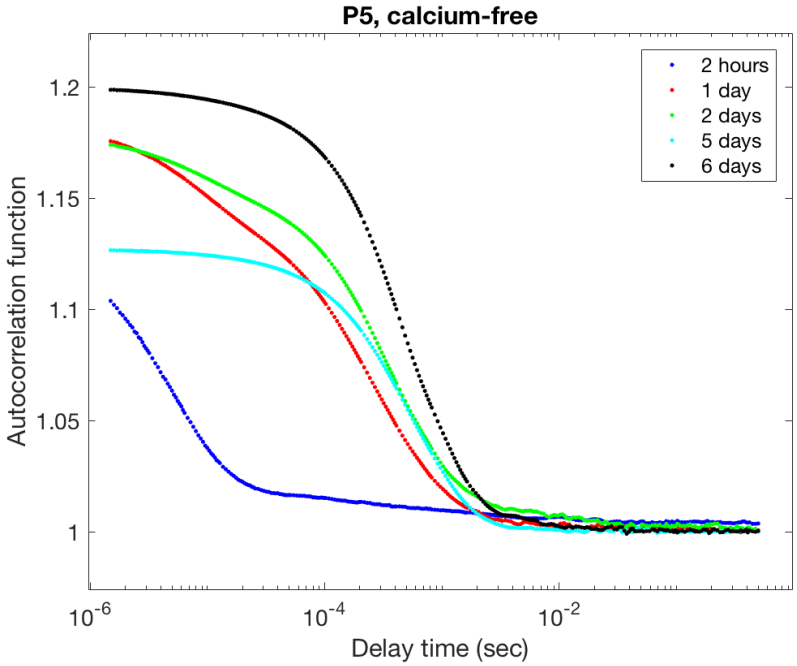

Figure 42: 2.5 mM P5 in Ca^{2+}-free buffer with 2 mM TRIS at pH 7.5, analysed using DLS. Non-normalized autocorrelation functions plotted against the delay time.

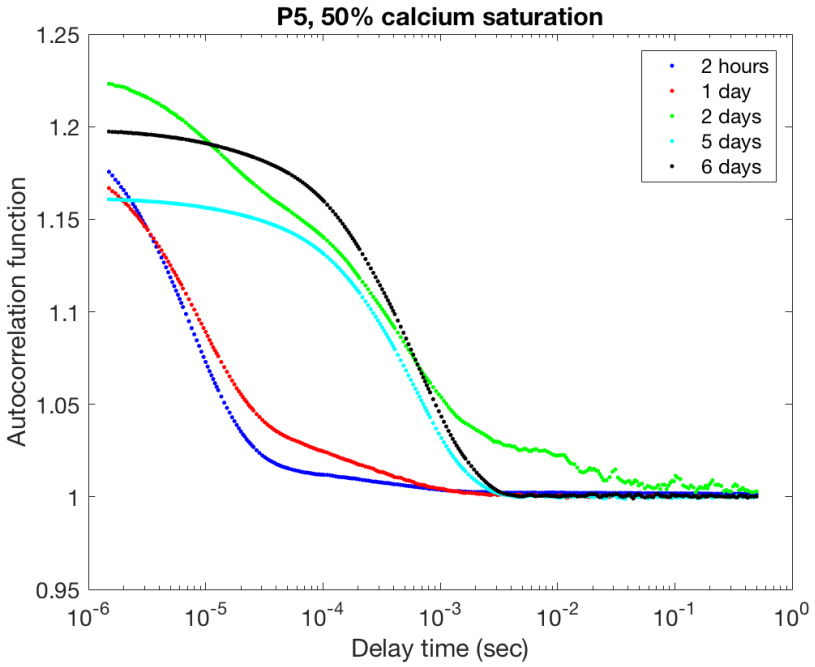

Figure 43: 2.5 mM P5 at 50% Ca^{2+} saturation with 2 mM TRIS at pH 7.5, analysed using DLS. Non-normalized autocorrelation functions plotted against the delay time.

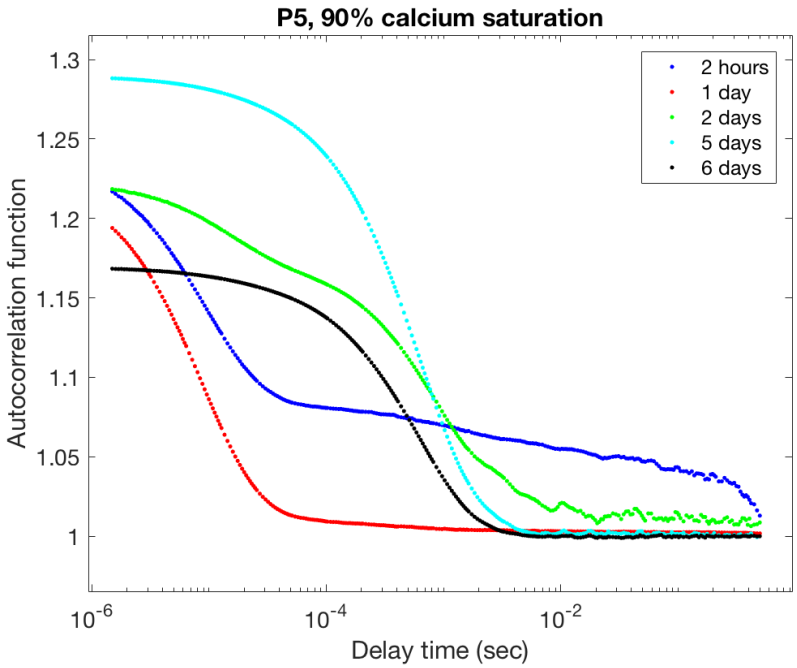

Figure 44: 2.5 mM P5 at 90% Ca^{2+} saturation with 2 mM TRIS at pH 7.5, analysed using DLS. Non-normalized autocorrelation functions plotted against the delay time.

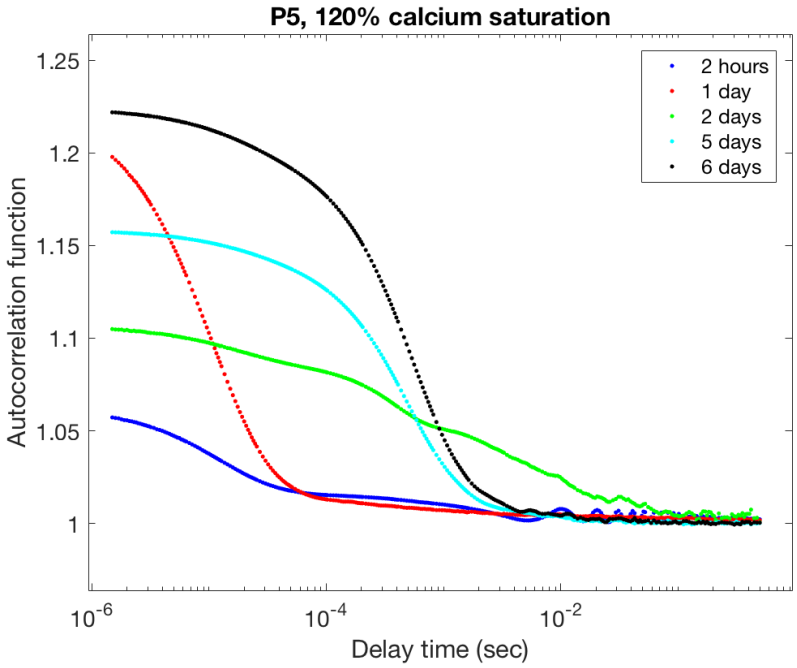

Figure 45: 2.5 mM P5 at 120% Ca^{2+} saturation with 2 mM TRIS at pH 7.5, analysed using DLS. Non-normalized autocorrelation functions plotted against the delay time.

Appendix C - PAGE

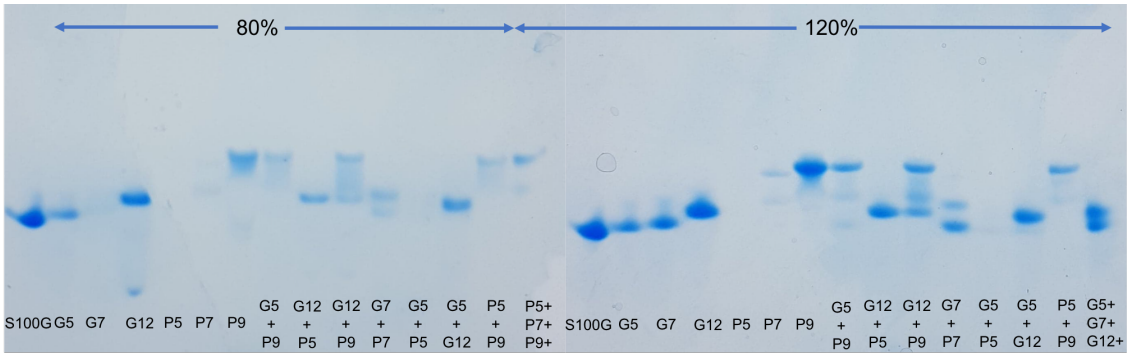

Figure 46: Native PAGE of various substrate samples with 1 mM protein in 2 mM TRIS buffer, pH 7.5 and incubated for two days at room temperature. 14% Tris-glycine gel and the staining method is Coomassie Brilliant Blue.

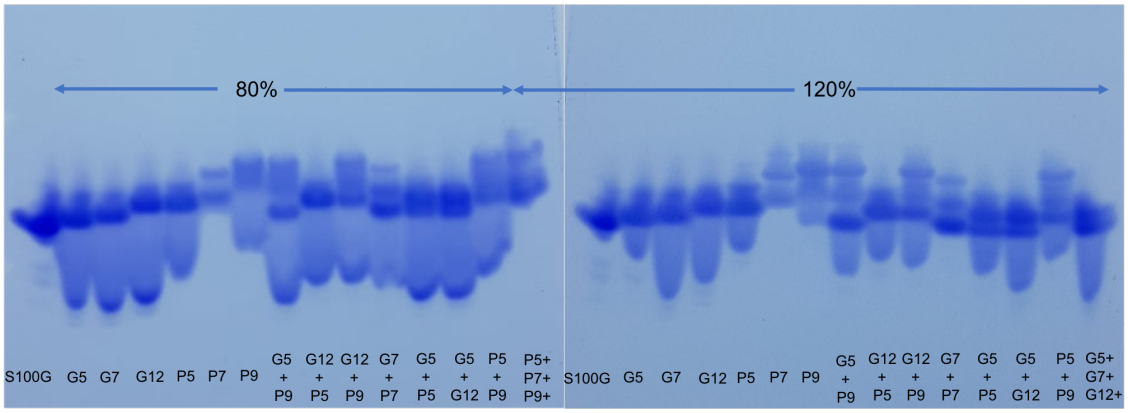

Figure 47: Native PAGE of various substrate samples with 1 mM protein in 2 mM TRIS buffer, pH 7.5 and incubated for four days at room temperature. 14% Tris-glycine gel and the staining method is Coomassie Brilliant Blue.

Appendix D - SEC

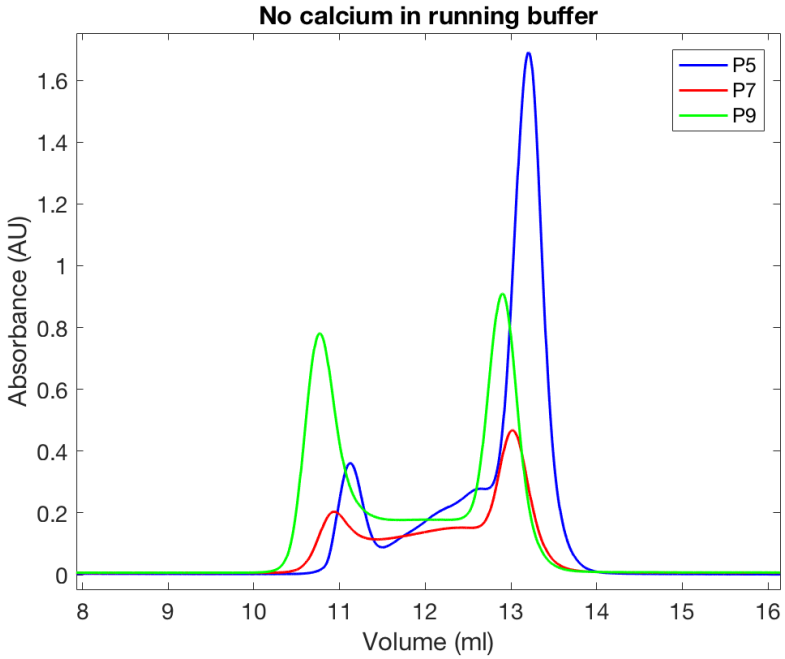

Figure 48: Non-normalized absorbance of the wavelength 214 nm, measured of P5, P7 and P9 eluates. The substrates were analysed using SEC directly after dissolving to 1 mM protein in 2 mM TRIS buffer, pH 7.5. No Ca^{2+} were added to the running buffer, but the samples had a saturation of 120% before injection.

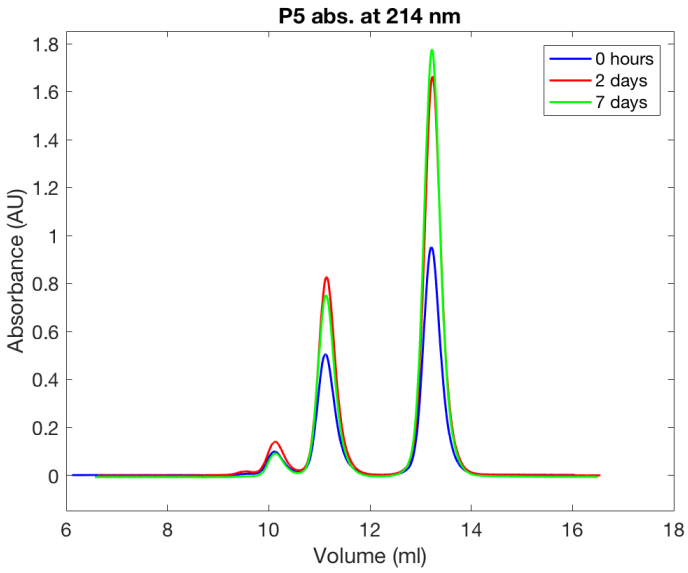

Figure 49: Non-normalized absorbance of the wavelength 214 nm of P5 eluate from SEC. The protein was incubated at the concentration 1 mM, in 2 mM HEPES buffer, pH 7.5, at 120% Ca^{2+} saturation and at room temperature. The measurements were done at the time points: directly after dissolution, after two days respectively seven days.

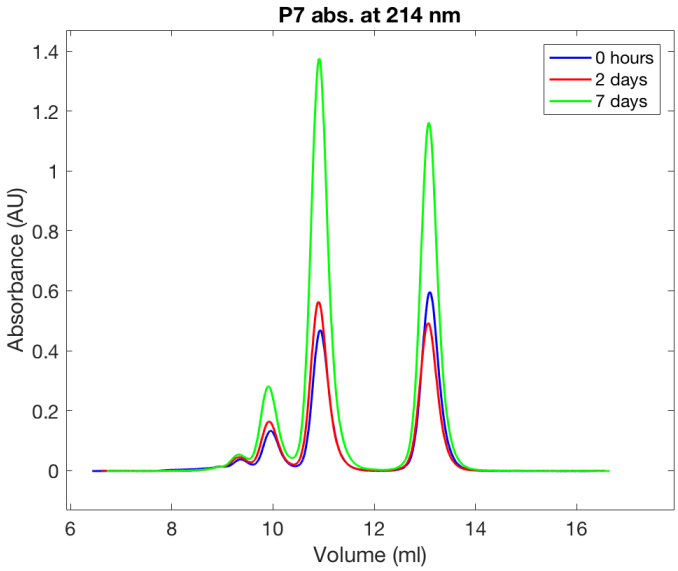

Figure 50: Non-normalized absorbance of the wavelength 214 nm of P7 eluate from SEC. The protein was incubated at the concentration 1 mM, in 2 mM HEPES buffer, pH 7.5, at 120% Ca^{2+} saturation and at room temperature. The measurements were done at the time points: directly after dissolution, after two days respectively seven days.

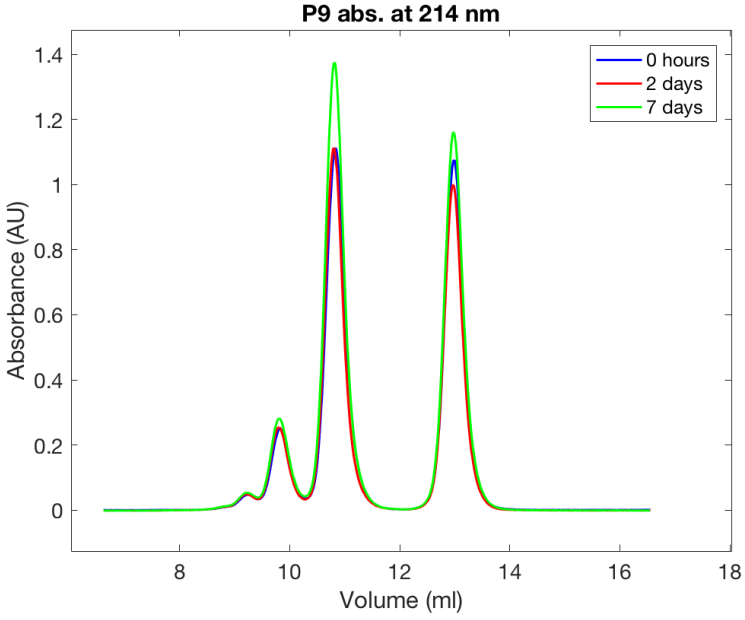

Figure 51: Non-normalized absorbance of the wavelength 214 nm of P9 eluate from SEC. The protein was incubated at the concentration 1 mM, in 2 mM HEPES buffer, pH 7.5, at 120% Ca^{2+} saturation and at room temperature. The measurements were done at the time points: directly after dissolution, after two days respectively seven days.

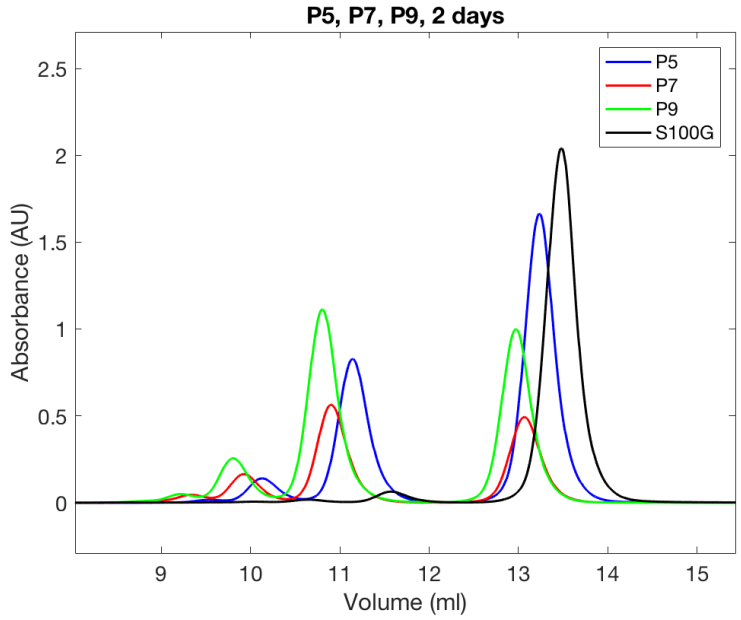

Figure 52: Non-normalized absorbance in the comparison of absorbance (214 nm light) of P5, P7 and P9. All of the proteins were incubated for 2 days at the concentration 1 mM, in 2 mM HEPES buffer, pH 7.5, at 120% Ca^{2+} saturation and at room temperature.

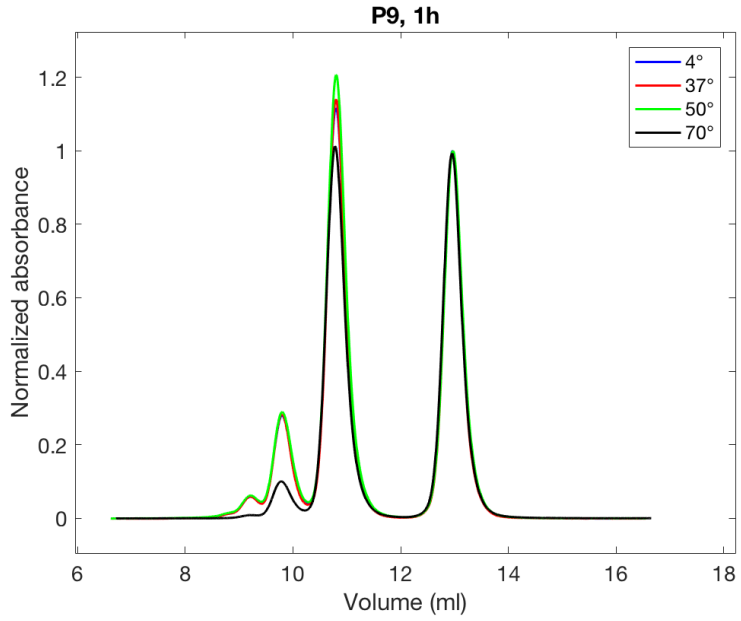

Figure 53: Absorbance (214 nm light) of eluate of P9 that had been incubated for one hour at various temperatures. Protein concentration was 1 mM, in 2 mM HEPES buffer, pH 7.5 and at 120% Ca^{2+} saturation. The absorbance are normalized with respect to the last peak. The blue and red spectrum overlap very much.

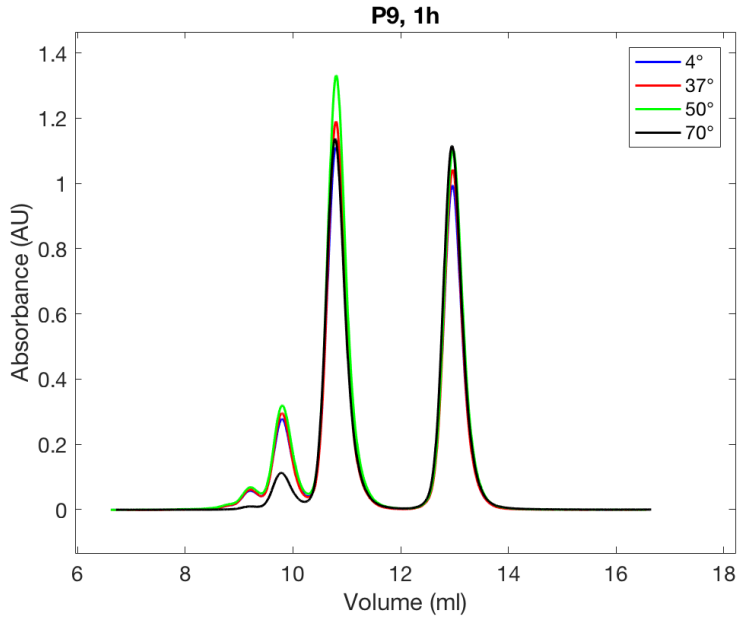

Figure 54: Non-normalized absorbance (214 nm light) of eluate of P9 that had been incubated for one hour at various temperatures. Protein concentration was 1 mM, in 2 mM HEPES buffer, pH 7.5 and at 120% Ca^{2+} saturation. The blue and red spectrum overlap very much.

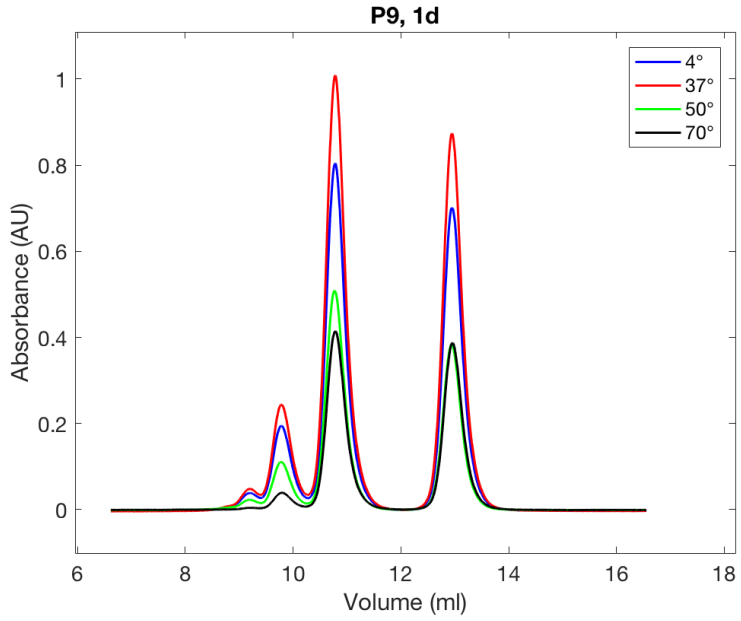

Figure 55: Non-normalized absorbance (214 nm light) of eluate of P9 that had been incubated for one day at various temperatures. Protein concentration was 1 mM, in 2 mM HEPES buffer, pH 7.5 and at 120% Ca^{2+} saturation.

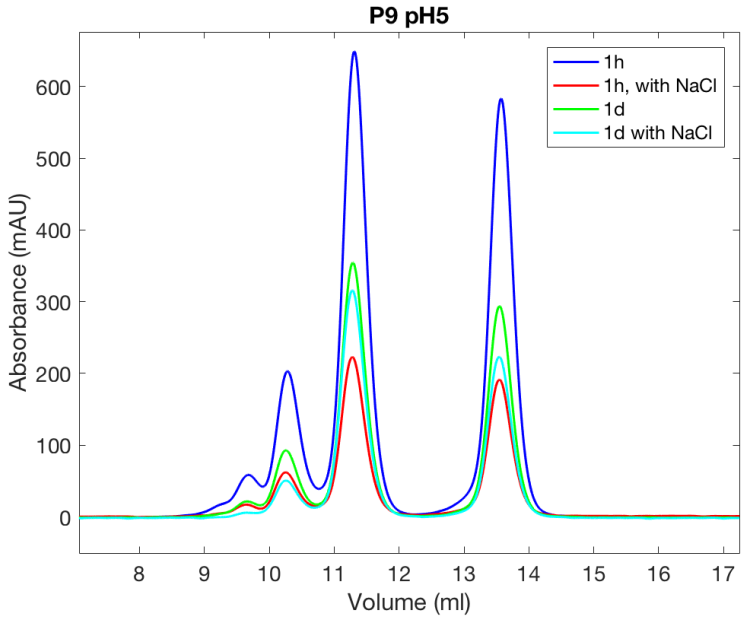

Figure 56: Non-normalized absorbance (214 nm light) of eluate of P9 at pH 5, with and without 0.15 M NaCl. Incubated for one hour respectively for one day at protein concentration 1 mM, in 2 mM MES monohydrate buffer, at 120% Ca^{2+} saturation and at 50°C.

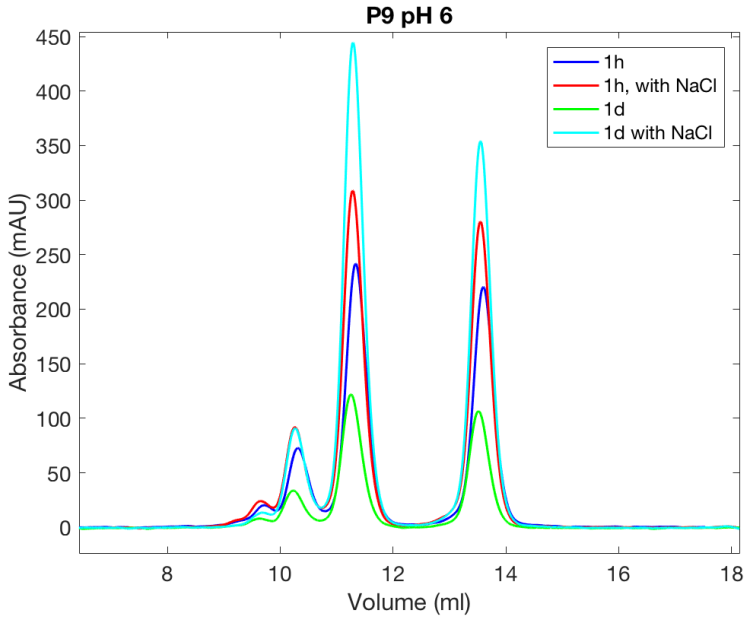

Figure 57: Non-normalized absorbance (214 nm light) of eluate of P9 at pH 6, with and without 0.15 M NaCl. Incubated for one hour respectively for one day at protein concentration 1 mM, in 2 mM MES monohydrate buffer, at 120% Ca^{2+} saturation and at 50°C.

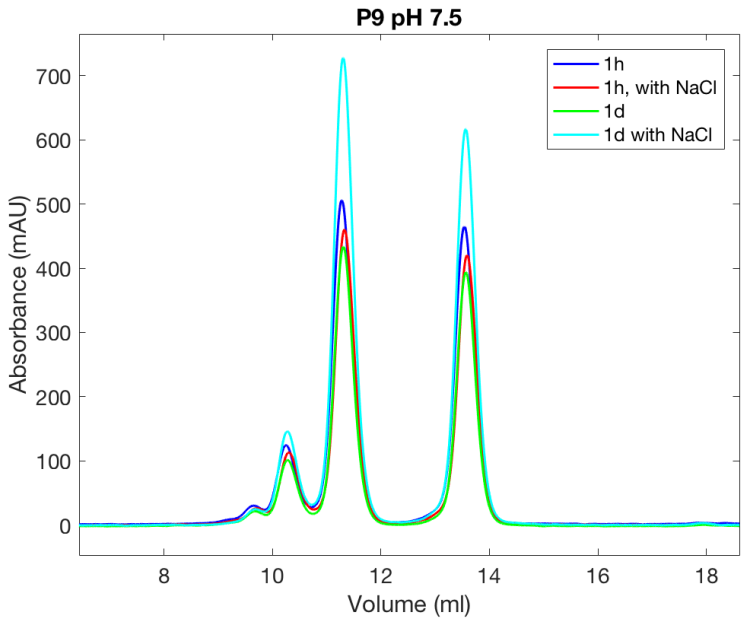

Figure 58: Non-normalized absorbance (214 nm light) of eluate of P9 at pH 7.5, with and without 0.15 M NaCl. Incubated for one hour respectively for one day at protein concentration 1 mM, in 2 mM HEPES buffer at 120% Ca^{2+} saturation and at 50°C.

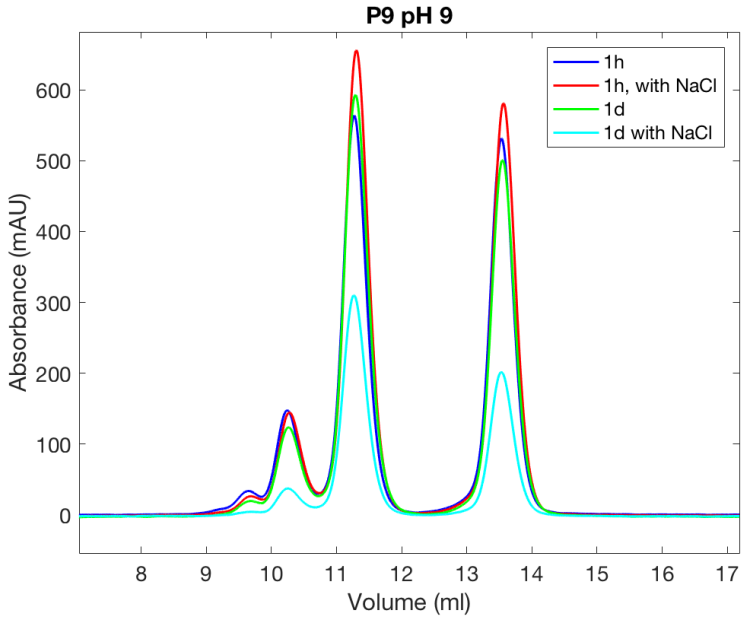

Figure 59: Non-normalized absorbance (214 nm light) of eluate of P9 at pH 9, with and without 0.15 M NaCl. Incubated for one hour respectively for one day at protein concentration 1 mM, in 2 mM Tricine buffer at 120% Ca^{2+} saturation and at 50°C.

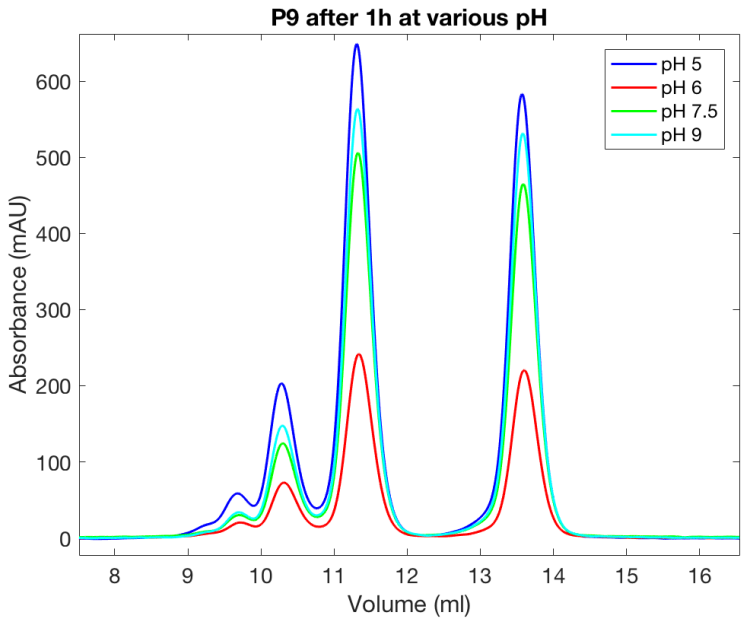

Figure 60: Non-normalized comparison of P9 at all four different pH, incubated for one hour at 50°C.

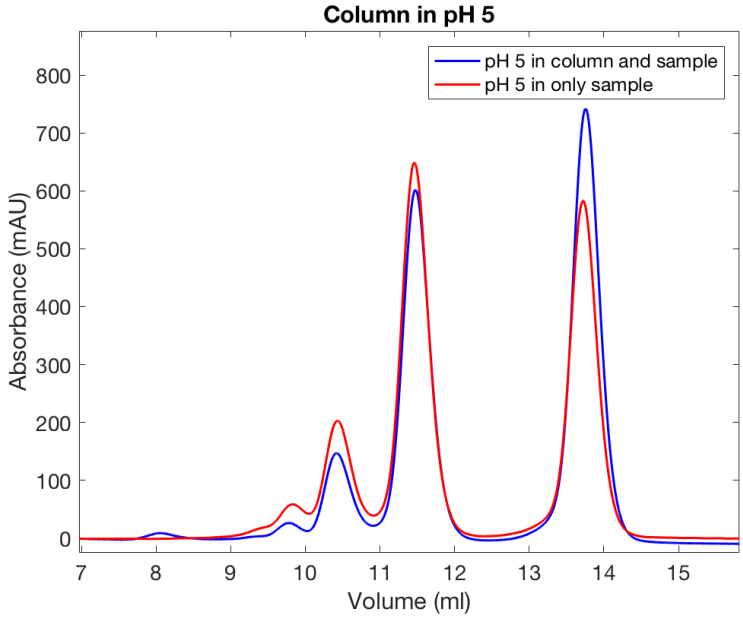

Figure 61: Non-normalized absorbance (214 nm light) of eluate of P9, incubated for one hour at protein concentration 1mM, in 2 mM MES monohydrate buffer, pH 5, at 120% Ca^{2+} saturation and at 50°C. The blue graph corresponds to the run where the running buffer had pH 5 and the red to the normal running buffer of pH 7.5.

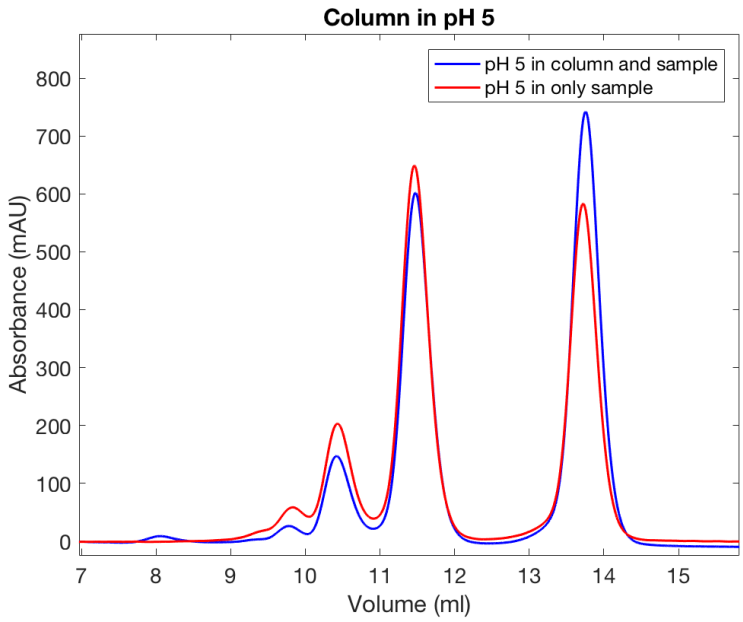

Figure 62: Non-normalized absorbance (214 nm light) of eluate of old and new P9.