

## 2 PROTEIN AGGREGATION IN BULK SOLUTION AND AT INTERFACES

Due to its potentially serious consequences protein aggregation has drawn major attention in recent years and has even been identified as a major obstacle to overcome in the development of protein formulations, for instance in the case of highly concentrated monoclonal antibody solutions<sup>22</sup>. As already denoted above, protein aggregation can occur at almost any stage of manufacturing, processing, storage, shipment and administration to the patient. During each of these steps, aggregation is governed by different influencing factors and critical parameters. For understanding and controlling aggregation during any of these steps it is necessary to experimentally isolate the factors triggering aggregation and studying them separately. Therefore, in the present work protein aggregation is not classified by structural characteristics as done in earlier works<sup>23</sup> such as type of bond, reversibility, size or conformation, but classified into the induction factors, that are causing protein aggregation. Since it is well-known that the aggregation behavior in bulk solution is fundamentally different from protein aggregation that involves adsorption to a bulk surface<sup>24</sup> these two phenomena are investigated separately throughout this thesis.

### 2.1 PROTEIN AGGREGATION IN BULK SOLUTION

Since protein aggregation is a critical phenomenon for the safety and efficacy of protein drugs, extensive research work has been dedicated to elucidate the factors controlling protein aggregation<sup>5-6,15,17,23,25-27</sup>. Although different proteins and a variety of influencing factors were studied there is some common ground between most of the investigations. The most common idea of how protein aggregation in solution proceeds is that partially unfolded states (also referred to as molten globule state or “A” states if the protein is acid-denatured) with reduced (but still substantial) secondary structure and clearly reduced tertiary structure expose hydrophobic surfaces and subsequently aggregate<sup>25</sup>. In order to suppress aggregation in bulk solution, it is necessary to maximize two physical protein properties: conformational and colloidal stability.

Increasing conformational stability means that the population of highly aggregation-prone partially unfolded intermediates has to be kept as low as possible. The relative degree of unfolding of the aggregation-prone intermediates is often very small (at most a few percent<sup>24</sup>) and spectroscopic techniques observing the overall, average conformation of a population of protein molecules might not be able to detect the subtle changes on the molecules. Nevertheless, aggregation in such solutions may rapidly proceed<sup>5</sup>. Oftentimes conditions that allow for a maximum conformational stability do not also provide the best conditions for maintaining the second parameter that should be maximized in order to reduce protein aggregation, colloidal stability. Therefore, often a compromise has to be struck in the

selection of the protein formulation. This compromise is still most conveniently achieved by empirical formulation studies.

Conformational stability can be increased by selecting favorable solution conditions. An important factor leading to favorable solution conditions is an appropriate solution pH. Many proteins tend to slightly unfold on a tertiary structural level when acidic conditions are chosen, such as IgG-antibodies or rh-GH, two proteins that are examined in this thesis. For example, a rhu-mAb anti-CD 20 antibody was found to lose its tertiary structure below pH 3<sup>14</sup>. Generally, weakly acidic conditions (pH 5-6) seem to be optimal for the formulation of mAbs<sup>6</sup>. Also rh-GH partially unfolds at low pH-values. For example, rh-GH is reported to populate a partially unfolded "A-state" at pH 2.5 that, in the presence of NaCl, leads to rapid aggregation of the protein<sup>28</sup>. In contrast, rh-GCSF maintains its conformational stability as determined by urea unfolding even at a low pH of 3.5<sup>29</sup>.

Another important factor that compromises conformational stability of proteins and therefore accelerates aggregation in bulk solution is temperature. High temperatures perturb the native protein conformation to a degree that accelerates aggregation<sup>17</sup>. Often aggregation starts well below the temperature that is experimentally determined as the melting temperature of the protein (the temperature at which 50 % of protein molecules are unfolded during a thermal transition<sup>23</sup>), validating the assumption that aggregates are not formed from fully unfolded monomers but that a certain fraction of partially unfolded monomer is sufficient to promote aggregation<sup>29</sup>. The thermal stability of proteins strongly varies. Compared to other proteins antibodies seem to be less sensitive to high temperatures taking into account their melting temperatures of above 70°C<sup>30</sup> whereas most other proteins already completely unfold below 70°C<sup>2</sup>. Increased aggregation rates upon temperature increase are also the basis of accelerated stability studies at elevated temperature carried out for the prediction of aggregation rates during the shelf-life of a protein. However, the assumption of a simple Arrhenius behavior allowing for the extrapolation of the accelerated stability (e.g. at 50°C) data to shelf-life data (e.g. at 4°C) can be seriously misleading<sup>31</sup> and was reported to potentially lead to the underestimation of the rate coefficient for monomer loss and hence to an overestimation of the shelf life of a therapeutic protein<sup>24,32</sup>. Nevertheless there is little alternative to that kind of studies since multi-year stability data at the target storage condition would not be available until late stages of clinical development at which any changes in the formulations would be very costly and difficult from a regulatory perspective.

Conformational stability can also be influenced by ligand binding. This is reflected by the Wyman linkage function which states that preferential binding of ligands to the native state of a protein is expected to shift the folding equilibrium towards a larger population of native protein molecules. Consequently the protein's propensity to aggregate will be reduced<sup>31,33-34</sup>. In contrast, preferential interaction with the unfolded or partially unfolded state of a protein

will result in a decrease of the thermal stability of the protein, e.g. as observed with the preservative benzyl alcohol when binding to interleukin 1 or rh-GCSF<sup>35-36</sup>. Ligand binding will be of special importance throughout this work since cyclodextrins are reported to preferentially bind to the unfolded state of proteins<sup>21</sup>, thereby potentially influencing conformational stability of the proteins under investigation. As discussed in further detail below, nonspecific stabilizing compounds like sucrose also influence conformational stability by being preferentially excluded from the protein surface.

Partial unfolding of protein molecules to highly aggregation-prone intermediates, as expressed by the conformational stability, is often the determining step in the formation of aggregates. However, also tendency of small aggregate nuclei to grow to larger aggregates can be a rate-limiting step in protein aggregation, generally referred to as colloidal stability. A global measure taking into account all sorts of intermolecular interactions between the protein molecules (van der Waals, electrostatic, hard-sphere) is given by the second virial coefficient ( $B_{22}$ ). Positive  $B_{22}$ -values indicate overall repulsive forces between the protein molecules in solution: protein-solvent interactions are favored over protein-protein interactions. In contrast negative  $B_{22}$ -values indicate attractive forces between protein molecules when protein-protein interaction is favored over protein-solvent interaction<sup>25</sup>. Since the  $B_{22}$ -value greatly depends on protein charge, alterations of the solution pH can have dramatic effects on the colloidal stability of a system. For rh-GCSF, one of the proteins that are investigated in this thesis, the role of colloidal stability is very well understood. At low pH (e.g. 3) the rh-GCSF molecules are positively charged and repulsive forces dominate. However, at neutral pH (between pH 5 and 7), aggregation rapidly proceeds, although conformational stability remains nearly unaltered, because the repulsive forces are no longer dominating. In addition to shifting the solution pH in a way that reduces repulsive forces between proteins, colloidal stability can also be lowered by the addition of salts leading to a shielding of repulsive forces<sup>25,29</sup>.

## 2.2 AGGREGATION AT THE AIR-WATER INTERFACE

In comparison to aggregation in bulk solution, the situation in the presence of large hydrophobic interfaces is fundamentally different, since new reaction pathways for protein aggregation are opened up. When partial unfolding of a protein is the rate limiting step for aggregation, the presence of an interface can massively increase aggregation rates. The reason for this phenomenon is that proteins are amphiphilic molecules and this property leads to their strong tendency to accumulate at interfaces. Most proteins exhibit a remarkable adsorption to hydrophobic surfaces, the air–water interface not only being among the most hydrophobic but also most frequently encountered interfaces, e.g. during mechanical agitation and mixing, spray-drying or filtration<sup>37</sup>. Layer thickness of the air-water interface is reported to be in the order of magnitude of about 2 nm which is about the same size as a

protein molecule<sup>5</sup> or 3.1 nm with a secondary layer below of about 5-7 nm thickness<sup>38</sup>. Other interfaces that therapeutic proteins are typically exposed to during their lifecycle may include the glass–water (in vials) or ice–water–interface (during freezing and thawing) which are discussed in the following section.

Generally protein adsorption to the air–water–interface can be divided into three steps<sup>39</sup>. First, diffusion of the protein–molecules into a subsurface has to take place. Proteins then have to overcome energy barriers (caused by surface pressure and an electrical bilayer) and adsorb to the surface. Finally proteins have to rearrange at the surface which involves partial unfolding of the adsorbed protein segments. By exposing parts of the hydrophobic protein core, contacts with the interface are maximized on both sides of the interface and the molecule regains conformational entropy<sup>39</sup>. A protein that is adsorbed to the air-water-interface experiences forces that are dramatically different from the forces in the bulk solution: it has been estimated that the tension forces perpendicular to the interface are as high as 140 pN and therefore large enough to unfold a protein<sup>5</sup>. The altered protein structure along with high local concentrations at the interface often lead to aggregation processes<sup>17</sup>. In addition, in agitated solutions a new air-water interface is continuously created thereby producing an amount of unfolded proteins that is no longer negligible compared to the amount of protein in the bulk and substantial aggregation often results. As discussed above, it is well-known that protein aggregation may have serious implications for the safety and efficacy of protein drugs<sup>6,40</sup>. Hence for a new protein formulation surface–induced aggregation during processing and storage has to be circumvented.

Agitation-induced aggregation has been reported<sup>17,41</sup> for a variety of proteins<sup>42-47</sup>, and it is a serious concern for the formulation of mAbs<sup>27,48-52</sup> and fusion proteins containing parts of immunoglobulins<sup>53</sup>. Aggregates formed by agitation have been determined to be very different in nature from aggregates of the same IgG-antibody formed during storage at elevated temperature<sup>50</sup>. Whereas insoluble heat-induced aggregates showed strong alterations of their secondary structure and did not redissolve into soluble aggregate components upon storage, insoluble aggregates formed by agitation-stress were demonstrated to maintain a very native-like conformation and to exist in equilibrium with other small aggregate types<sup>50</sup>. The degree of mAb-aggregation after agitation is influenced by a variety of parameters. The first parameter is the structure of the mAb itself since some IgG antibodies are reported to significantly aggregate within hours of agitation<sup>27,52</sup> whereas others are reported to exhibit a remarkable resistance to aggregation at the air-water interface, after two weeks of agitation at 200rpm<sup>54</sup> or even after two weeks of shaking in vials<sup>51</sup>. It has been suggested that for the successful development of monoclonal antibodies the surface activity of the potential drug candidate should be taken into account, since a positive correlation of susceptibility to shaking-induced aggregation and surface-activity was

reported<sup>55</sup>. However, it seems that also significantly surface-active mAbs can be very resistant to agitation-induced aggregation<sup>51</sup>. Furthermore the degree of IgG-aggregation during agitation-studies is strongly influenced by the filling volume and the existence of a head space in the shaken container vial<sup>27</sup>. In the absence of a head-space (exchange of the air-water-interface by a glass-water interface) the IgG-antibody remains stable whereas the existence of a head space causes significant aggregation. Finally also the concentration as well as the type of ions is found to have an influence on mAb-aggregation during agitation<sup>52</sup>. With increasing ionic strength agitation-induced aggregation is increased. The nature of the examined cations does not influence aggregation, however the selection of anions has a strong influence on shaking-induced aggregation<sup>52</sup>.

Also for recombinant human growth hormone (rh-GH) extensive investigations were carried out regarding the behavior after mechanical stressing and exposure to the air–water–interface. Rh-GH was found to aggregate after vortexing or when being shaken in glass vials<sup>43,47</sup>. In addition investigations were carried out that concluded that shear forces alone cannot be made responsible for rh-GH–aggregation after mechanical stressing but that the presence of an air–water–interface is a necessary prerequisite<sup>56</sup>. This behavior was demonstrated by investigations using a rotor–stator–device and a nitrogen–bubbling–method<sup>57</sup>. Furthermore rh-GH tends to aggregate in the presence of other hydrophobic surfaces such as PTFE whereas the behavior under thermally denaturing conditions cannot be correlated to denaturation at hydrophobic surfaces<sup>58</sup>. Similar findings are reported for lysozyme and insulin inactivation in the presence of the hydrophobic surfaces PTFE and air whereas the presence of less hydrophobic glass material caused a smaller degree of inactivation<sup>46,59</sup>.

Little studies are available on the behavior of the third model protein of this thesis, rh-GCSF, during agitation. In studies on PEG-GCSF it was found that there is an inverse relationship between concentration of the protein and susceptibility to agitation-induced aggregation<sup>53</sup>. Since a later work has found that the aggregation mechanism of PEG-GCSF is very similar to that of rh-GCSF it can be assumed that the findings for PEG-GCSF apply to rh-GCSF in a similar manner<sup>60</sup>.

### **2.3 AGGREGATION DURING FREEZE-THAWING AND IN THE PRESENCE OF MICROPARTICLES/VARIOUS SURFACES**

As a third major induction factor for protein aggregation, freezing and thawing (F/T) processes are discussed. F/T processes occur at multiple stages during manufacturing, processing, storage and analytics of protein pharmaceuticals<sup>17</sup>. For instance, protein bulk solutions are routinely stored at -70°C as an intermediate step during commercial protein pharmaceuticals production, assuming increased long-term stability as compared to storage in the liquid state. For subsequent processing bulk solutions have to be thawed again.

Protein solutions may also be unintentionally frozen due to inappropriate handling of the final parenteral protein products and finally protein samples may also be frozen and thawed later for analytical purposes when analytics cannot be carried out immediately. All of the mentioned processes may also occur repeatedly, thereby exposing the proteins to significant stress that has to be overcome<sup>17</sup>.

Numerous studies identifying factors influencing protein stability during freezing and thawing and characterizing resulting protein instability are available. Obviously, the factors controlling protein stability in solution - conformational and colloidal stability - also influence a protein's susceptibility to freeze-thawing-induced degradation with pH and ionic strength being the key parameters<sup>29,61-62</sup>. In addition, some further factors specific for freeze-thawing-induced stress also influence the extent and the characteristics of protein instability. It was found that freezing by itself can perturb a protein's native conformation: cold denaturation<sup>63</sup>. Freezing processes can also lead to freeze-concentration-processes with locally increased protein-concentration that can result in elevated aggregation rates as already discussed above along with the section dealing with aggregation at the air-water interface. In addition, exposure to the ice-water-interface is reported to induce protein unfolding and subsequent aggregation processes, rendering freeze-thawing-stress a further surface-induced protein instability<sup>58</sup>. Since exposure to the ice-water-interface triggers protein instability, it has to be assumed, that protein-concentration is of importance, because a more favorable protein-surface-ratio can be achieved at high protein concentrations thereby decreasing the rate of aggregation. A lower fraction of protein exposed to the surface also explains why there are several reports on decreased protein aggregation despite increasing protein concentration<sup>62</sup>, which usually leads to accelerated aggregation rates in solution as experienced with highly concentrated antibody formulations<sup>22,64</sup>. However, it is reported that this rule does not necessarily always hold true for antibodies, because it is reported that the increase of the concentration of a chimeric antibody (L6) did not inhibit F/T-induced aggregation<sup>14,54</sup>.

Exposure of the protein to the ice-water-interface also explains why the freeze-thawing-rates can have an influence on protein stability. One would expect that very fast freezing- and thawing-rates minimize damage of the proteins because that way the time of exposure to the harmful ice-water interface is as short as possible. However, several reports state that even very fast freezing and thawing, for instance achieved by immersion into liquid nitrogen, did not stabilize the proteins under investigation compared to slower freezing and thawing rates<sup>58</sup>. In contrast, too slow freezing rates may foster crystallization of solution components thereby leading to accelerated aggregation rates<sup>6,65</sup>. Finally, also the container material and geometry as well as its size can be critical for protein stability, since they also alter warming and cooling rates and the extent of exposure to the ice-water-interfaces as well as to the container-liquid interface<sup>62</sup>. Consequently the prediction of freeze-thawing-induced

aggregation of large bulk quantities from small-scale stress testing poses a major challenge and if availability of protein material allows it, freeze-thawing-induced damage of the protein should always be evaluated at scale<sup>23</sup>. Also, since during thawing of large bulk quantities of protein containers are usually shaken, thawing steps include mechanical stress of the protein.

For the sake of completeness it should be mentioned that also further surfaces are capable of accelerating protein aggregation. For instance, silica microparticles can be shed from glass vials during the autoclaving procedure<sup>5</sup> and such microparticles were shown to induce heterogeneous nucleation processes in recombinant human platelet-activating factor acetylhydrolase, which lead to significant aggregation<sup>26</sup>. Removing the exogenous particles by filtration suppressed aggregation processes. Another study did not observe increased mAb-aggregation in the presence of glass-microparticles but nevertheless suggested using a testing protocol to routinely examine the potential effects of micro- and nanoparticles that could be shed from wetted surfaces<sup>66</sup>. Further solid-liquid interfaces that therapeutic proteins can be exposed to during their lifecycle and that were demonstrated to potentially compromise protein stability include the Teflon™-water interface (aggregation of insulin)<sup>67</sup>, stainless steel particles shed from a filling piston pump causing mAb-aggregation at their interface with the mAb-solution<sup>68</sup>. Also leachates from tungsten as well as silicone oil syringe lubricant were already reported to cause protein precipitation<sup>69-70</sup>.