Aggregation of Recombinant Human Granulocyte-Colony Stimulating Factor (rhG-CSF) under Process Conditions

Ulrich Roessl, BSc

Aggregation of Recombinant Human Granulocyte-Colony Stimulating Factor (rhG-CSF) under Process Conditions

Master's Thesis

at

Graz University of Technology

submitted by

Ulrich Roessl, BSc

Research Center Pharmaceutical Engineering GmbH
Institute for Biotechnology and Bioprocess Engineering
Graz University of Technology
A-8010 Graz, Austria

December 12th, 2010

Supervisor: Univ.-Prof. Dipl.-Ing. Dr.-techn. Bernd Nidetzky

TU Graz

Abstract

Just like many other protein therapeutics recombinant human granulocyte-colony stimulating factor (rhG-CSF) tends to form soluble and insoluble aggregates during storage. Aggregation massively impairs a drug's shelf life and therefore represents probably the most troubling manifestation of protein instability. Little is known about harsh process conditions as mechanical stress and surface effects and their impact on protein aggregation. In this study we investigated the influence of such process conditions on aggregation kinetics and the underlying mechanisms. Process kinetic analysis and fluorescence spectroscopy revealed that conformational perturbance is an essential factor for rhG-CSF aggregation. The formation of new, partially intermolecular disulfide bonds was shown with mass spectrometry. The single free cysteine in native rhG-CSF was found to be highly involved in disulfide rearrangement. Confocal laser scanning microscopy facilitated comparison of size and shape of insoluble aggregate particles. The exposure of rhG-CSF to air-liquid interfaces induced both aggregation and structural expansion more effectively than agitation under the investigated conditions. Minimization of surface exposure during storage and production would very likely lead to a reduction of aggregate formation.

Kurzfassung

So wie viele andere pharmazeutische Proteine neigt rekombinanter humaner Granulozyten-Kolonie stimulierender Faktor (rhG-CSF) zur Bildung von löslichen und unlöslichen Aggregaten während längerfristiger Lagerung. Aggregation beeinträchtigt die Haltbarkeit solcher Therapeutika massiv und stellt daher besondere Herausforderungen an die Formulierung. Über die Auswirkungen von rauen Prozessbedingungen wie mechanischem Stress und Oberflächeneffekten auf Proteinaggregation, sowie deren zugrunde liegende Mechanismen ist wenig bekannt. Mithilfe von prozesskinetischer Analytik und Fluoreszenzspektroskopie konnte in dieser Arbeit gezeigt werden, dass strukturelle Konformationsänderungen elementarer Bestandteil der Aggregation von rhG-CSF sind. Massenspektrometrie bestätigte die Neuformierung von teils intermolekularen Disulfidbrücken. Konfokale Laser Scanning Mikroskopie ermöglichte den Vergleich von Größe und Form unlöslicher Aggregatpartikel. Gas-flüssig-Oberflächeneffekte bewirkten stärkere strukturelle Konformationsänderungen sowie erhöhte Aggregation als Rühren unter den untersuchten Bedingungen. Die Minimierung von Oberflächeneffekten während Lagerung und Produktion würde mit hoher Wahrscheinlichkeit zu einer Verminderung der Proteinaggregation führen.

Statutory Declaration

I declare that I have authored this thesis independently, that I have not used other than the declared sources / resources, and that I have explicitly marked all material which has been quoted either literally or by content from the used sources.

Place	Date	Signature

Eidesstattliche Erklärung

Ich erkläre an Eides statt, dass ich die vorliegende Arbeit selbstständig verfasst, andere als die angegebenen Quellen/Hilfsmittel nicht benutzt, und die den benutzten Quellen wörtlich und inhaltlich entnommene Stellen als solche kenntlich gemacht habe.

Ort	Datum	Unterschrift

For Magdalena

Contents

Abstract **III**

Statutory Declaration **VI**

Contents **IX**

Acknowledgements **X**

1 Introduction **1**

2 Materials and Methods **5**

2.1 Stressing Methods 5

2.2 Fluorescence Spectroscopy 6

2.3 Size Exclusion HPLC 6

2.4 Mass Spectrometry 6

2.5 Confocal Laser Scanning Microscopy 8

2.6 Kinetics and Statistics 8

2.7 Materials 8

3 Results **9**

3.1 Protein Loss 9

3.2 Conformational distortion probed by fluorescence spectroscopy 11

3.3 Determination of soluble aggregate fractions 11

3.4 Involvement of disulfide bridges 12

3.5 Size and shape of insoluble aggregates 16

4 Discussion **18**

4.1 Aggregation and Precipitation . . . 18

4.2 Exposure of hydrophobic domains . . . 19

4.3 Disulfide Rearrangement . . . 19

4.4 Insoluble Particle Distributions . . . 20

4.5 Aggregation Kinetics . . . 21

4.6 Conclusions . . . 23

5 Appendix **25**

5.1 Stirring . . . 25

5.2 Aeration . . . 26

5.3 UV Spectrophotometry . . . 27

5.4 Fluorescence Spectroscopy . . . 28

5.5 Size exclusion HPLC . . . 29

5.6 Dialysis and thin layer chromatography (TLC) . . . 29

5.7 SDS-PAGE . . . 30

5.8 Mass spectrometry . . . 31

5.9 Confocal Laser Scanning Microscopy . . . 31

5.10 Materials . . . 32

List of Figures **33**

Bibliography **37**

Supplementary Information **38**

Supplementary Information A . . . 39

Supplementary Information B . . . 43

Supplementary Information C . . . 47

Supplementary Information D . . . 51

Acknowledgements

First of all I would like to thank Johanna Wiesbauer. She guided me through this thesis with infinite patience and plenty of precious advice. I am deeply grateful for her permanent support and understanding.

I am also very indebted to Prof. Bernd Nidetzky, who gave me the opportunity to work in the field of protein technology. I am very proud of being part of his excellent research and to benefit from his remarkable experience.

Thanks to Stefan Leitgeb for unconfined help with all my problems, professional or private.

Gratefully, I want to acknowledge the help of Massimiliano Cardinale, Institute of Environmental Biotechnology, University of Technology Graz, for all the time he sacrificed for familiarizing me with confocal laser scanning microscopy.

Very special thanks to Ruth Birner-Grünberger for providing me with her great MS expertise, for all the labor she had with the analysis and particularly for her sympathy for me, being such an MS-rookie.

Thank you, Maggie, for your help and tolerance with me throughout the year.

Finally, I would like to express my deep gratitude to my parents. Without their unconditional support this thesis would never have been possible.

Chapter 1

Introduction

Protein aggregates are known to be responsible for numerous neurological disorders, such as Alzheimer's, Parkinson's and Huntington's disease. Researchers worldwide are trying to elucidate mechanisms involved in protein aggregation to enable treatment and prophylaxis [1]. Besides that field, protein aggregation is a matter of big concern in pharmaceutical industry [2]. As a major factor determining a drug's shelf life protein aggregation is target of enormous efforts in order to optimize drug formulations [3][4]. If administered parenterally, protein aggregates can cause reduced pharmaceutical effectiveness but also severe immunological reactions up to anaphylactic shock [5][6]. Minimizing protein aggregation in formulations today is an empiric optimization process which is individual for every protein and thus time consuming and expensive. There are several in silico-approaches to predict a protein's aggregation tendency, e.g. depending on its primary structure [7][8]. Most of them focus primarily on amyloid β-aggregation which is the predominant form in neurodegenerative human diseases. Other types of aggregation, for instance formation of disulfide bonds, remain disregarded . Moreover, the respective proteins, e.g. amyloid β-peptide, are usually shorter than many therapeutic proteins, such as cytokines or antibodies.

To provide a basis for the establishment of innovative in-silico approaches a detailed understanding of aggregation mechanisms and process kinetics is crucial. Protein aggregation can occur in virtually any step of biopharmaceutical production. It can be induced by temperature, freezing/thawing, shaking, stirring, pumping and other mechanical stress. Also surface effects are known to promote aggregation, both on liquid-solid and liquid-gas interfaces [3]. Only little is known about aggregation mechanisms and process kinetics under these conditions.

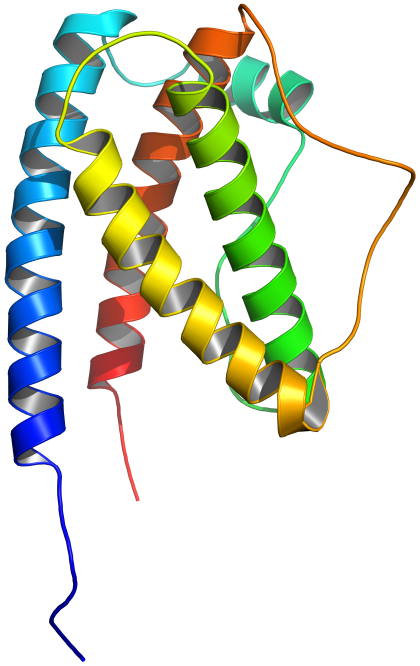


Figure 1.1: NMR structure of human G-CSF [11]

We chose recombinant human granulocyte-colony stimulating factor (rhG-CSF) from *E.coli* (Filgrastim) as a model protein to study aggregation under process conditions. rhG-CSF is a cytokine produced mainly in macrophages and shows high sequence similarity with members of the interleukin-6 superfamily. It induces proliferation of neutrophil colonies and differentiation of precursor cells to neutrophils, and it stimulates the activity of mature neutrophils [9].

As Filgrastim and Lenograstim it is therapeutically applied in oncology and hematology [10]. rhG-CSFs crystal structure has been determined in 1993 [9]. It exhibits a four-α-helix bundle motif with up-up-down-down connectivity, which is rather common among cytokines (Figure 1.1). In contrast to natural human G-CSF, Filgrastim exhibits methionine as primary amino acid and lacks three residues at respective positions 37-39. If not otherwise indicated, we are referring to that isoform whenever rhG-CSF is mentioned.

rhG-CSF is known to aggregate easily upon stressing and storage. Optimization of rhG-CSF stability in pharmaceutical formulations is of utmost interest for the producers. Within interleukin-6 superfamily, members exhibit very different aggregation behaviors. Preliminary experiments proved sizable differences in sensitivity for stressing factors and aggregation kinetics compared to rhG-CSF, despite the high structural similarity. Characterization of rhG-CSF

aggregation behavior under process conditions would be a first step towards understanding the observed differences.

rhG-CSF exhibits five cysteine residues. Only C^{18} does not form a disulfide bond in native state and is thus potentially susceptible to reduction. In fact, mutation to alanine decreased aggregation kinetics but did not inhibit it [12]. A partial involvement of disulfide bridges is thus likely and has even been investigated using SDS-PAGE [13]. However, there is no evidence so far that C^{18} is actually responsible for disulfide aggregation in rhG-CSF.

Filgrastim is produced in *E.coli* as inclusion bodies and lacks glycosylation at T^{134}. This seems to promote aggregation compared to the glycosylated Lenograstim, produced in CHO cells [14] [10]. Speculations on interference of the glycosylated loop region with C^{18} were largely disproved by a NMR study on the effects of glycosylation in rhG-CSF [15]. A conformational stabilization through the glycosylation seems more likely from the results [16]. This supports the hypothesis that a certain structural perturbation is prerequisite for rhG-CSF aggregation.

Chi and coworkers have proposed a general reaction mechanism for rhG-CSF [17]. It suggests that aggregation occurred via a structurally expanded transition state species of the monomer, which forms due to increased energy. Irreversible dimerization was only possible with two of the activated monomer species. Based on that model we tried to mimic process conditions to apply the activation free energy and hence induce rhG-CSF aggregation.

As mentioned above, production processes consist of numerous complex unit processes. Each of them is a result of several biophysical phenomena. For instance, it is impossible to apply mechanical stress without the comprehension of any surface effects or vice versa. Minimal variations of temperature due to stressing cannot be prevented, either. It has been reported before that shear stress and air-liquid interfaces promote protein denaturation and aggregation [18]. However, to look at the effects of those stressing factors separately is hardly feasible. Our approach for that problem was to create conditions where one stressing factor is clearly dominating over the others. We decided to apply mechanical stress by stirring, to expose the protein to air-liquid-interfaces by bubble aeration and to simply incubate rhG-CSF at 37°C.

We chose stirring speed and aeration rate in a way that considerable aggregation was observable through loss of protein and turbidity increment. As stirring occurred with a magnetic

stir bar a hydrodynamic description of the system or determination of energy input was not possible. Exposure to air-liquid-interfaces occurred via bubble aeration. Due to foam formation, the exact interface area could not be defined, either. Thus, we did not pursue quantification of the dependence of aggregation on distinct stressing factors. We rather tried to emphasize the differences in kinetics and aggregation mechanism between the stressing methods.

For monitoring of protein aggregation we observed the concentration of soluble rhG-CSF. Intrinsic fluorescence and fluorescence with 8-anilino-1-naphthalenesulfonic acid (ANS) should report structural changes during aggregation processes indirectly. ANS exhibits increased fluorescence activity in apolar environments [19]. Interaction with hydrophobic protein patches would lead to increased fluorescence. Quantification of soluble aggregates was accomplished with size exclusion high performance liquid chromatography (SEC). Confocal laser scanning microscopy (CLSM) enabled studying the changes in size and shape of insoluble protein aggregates. Size measurements of soluble aggregates, e.g. with dynamic light scattering, was not employed. The very broad size distribution of aggregate particles would have necessitated the operation of several different devices. Molecular forces involved in aggregation, especially disulfide bridges, were investigated with sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and protein sequence analysis with mass spectrometry (MS).

Chapter 2

Materials and Methods

2.1 Stressing Methods

Agitation and aeration were performed in double walled miniature bioreactors with a diameter of 32 mm. Temperature control occurred with a Julabo F25 Refrigerated/Heating Circulator. For stirring a magnetic stir bar (30x6 mm) was agitated at 400 rpm for 47 hours in 15 ml rhG-CSF solution. Six stirring experiments were carried out. For twelve hours of bubble aeration compressed air was fed into 10 ml rhG-CSF solution via a Stasto pressure reducer (R-M14-08-R) followed by a flow gauge to allow a constant flow rate of 3 liters per hour. The air flow was piped directly into solution with an autoclavable plastic tube LAB/FDA/USP grade VI with 3.2 mm inner diameter. The aeration experiment was repeated five times. Considerable solvent evaporation was observed during aeration which was minimized through pre-humidification of the air stream.

Protein concentration was adjusted to 1.39 ± 0.03 mg/ml. After centrifugation at 18,400 g to remove the precipitate, protein concentration in the supernatant was in all cases determined by measurement of spectroscopic absorbance at 280 nm with a molar extinction coefficient of 15,720 M^{-1} cm^{-1} (Beckman Coulter DU 800 spectrophotometer).
Turbidity was inspected visually.

All stressing experiments were performed in 10 mM glutamate buffer, pH 4.4, containing 5 % (w/v) sorbitol. Removal of sorbitol from the rhG-CSF bulk solution, as obtained from Sandoz, was accomplished with microdialysis through a Spectra/Por 4 membrane, MWCO 12-14,000. Dialysis performance was checked with thin layer chromatography on silica gel F254 aluminum sheets from Merck, and $KMnO_4$/NaOH stained.

2.2 Fluorescence Spectroscopy

Fluorescence measurements were performed on a Hitachi F-4500 fluorescence spectrophotometer equipped with a Julabo F25 refrigerated/heating circulator at 25°C, consistently. Intrinsic fluorescence was read at 340 nm from excitation wavelength of 280 and 295 nm, respectively. rhG-CSF concentration was adjusted to 92 μg/ml. ANS fluorescence was induced at 388 nm and read at 470 nm. 10 μl of 1 mM ANS was added to 500 μl rhG-CSF (92 μg/ml). Molecular ANS/rhG-CSF ratio was 4:1. Slit width was 5 nm in all fluorescence measurements.

2.3 Size Exclusion HPLC

SEC was carried out on a Merck Hitachi LaChrom system with an L-7480 fluorescence detector. Excitation occurred at 280 nm, emission at 345 nm. As mobile phase we used 50 mM $(NH_4)HCO_3$, pH 7.0 for chromatography over a TSK-GEL G3000SWXL column with precolumn from Tosoh Bioscience.

2.4 Mass Spectrometry

Samples for mass spectrometry were stirred for 22.5 hours or aerated for 15.5 hours, respectively. Precipitate from 5 ml remaining solution was concentrated by one minute centrifugation at 1,500 g and decanting of the supernatant. After resuspension in the remaining supernatant 10 μl were incubated for 30 minutes with 10 μl of 20 mM dithiothreitol (DTT) at 56°C for reduction of disulfide bonds. For alkylation of the free cysteines, 10 μl of 165 mM iodoacetamide (IAA) were added prior to incubation for 30 minutes at 37°C. Another 10 μl of rhG-CSF precipitate solution were treated only with 10 μl IAA and 10 μl highly pure water instead of DTT. Estimated 30 μg stirred rhG-CSF was applied for SDS-PAGE per lane. Only a fractional amount actually entered the gel due to the high amount of high molecular weight aggregates and precipitates.

Sample excision from the gel, enzymatic digestion, sample preparation and MS analysis were performed at the Center for Medical Research, Medical University Graz. Protein identification and internal sequence information was received from LC-MS/MS. Protein bands stained with Coomassie Brillant Blue R350 were excised from SDS gels and digested with Promega modified trypsin according to the method of Shevchenko et al. [20], or with 0,5 μg

chymotrypsin (Roche) in 50 mM ammonium bicarbonate and 10 mM $CaCl_2$. Digests were separated by nano-HPLC (Agilent 1200 system, Vienna, Austria) equipped with a Zorbax 300SB-C18 enrichment column (5 μm, 5 x 0.3 mm) and a Zorbax 300SB-C18 nanocolumn (3.5 μm, 150 x 0.075 mm). 40 μl of sample were injected and concentrated on the enrichment column for 6 min using 0.1 % formic acid as isocratic solvent at a flow rate of 20 μl/min. The column was then switched in the nanoflow circuit, and the sample was loaded on the nanocolumn at a flow rate of 300 nl/min. Separation was carried out using the following gradient, where solvent A is 0.3 % formic acid in water and solvent B is a mixture of acetonitrile and water (4:1, v/v) containing 0.3 % formic acid: 0-6 min: 13 % B; 6-35 min: 13-28 % B; 35-47 min: 28-50 % B, 47-48 min: 50-100 % B; 48-58 min: 100 % B; 58-59 min: 100-13 % B; 59-70 min: re-equilibration at 13 % B. The sample was ionized in the nanospray source equipped with nanospray tips (PicoTipTM Stock # FS360-75-15-D-20, Coating: 1P-4P, 15 $\pm$ 1 μm Emitter, New Objective, Woburn, MA, USA). It was analyzed in a Thermo LTQ-FT mass spectrometer (Thermo Fisher Scientific, Waltham, MA, USA) operated in positive ion mode, applying alternating full scan MS (m/z 400 to 2000) in the ion cyclotron and MS/MS by collision induced dissociation of the 5 most intense peaks in the ion trap with dynamic exclusion enabled. The LC-MS/MS data were analyzed by searching the SwissProt public database downloaded on May 26th, 2010, and the amino acid sequence of the cloned rhG-CSF for detailed modification analysis, with Mascot 2.2 (MatrixScience, London, UK). Search criteria were charge 2+ or 3+, precursor mass error 0.05 Da and product mass error 0.7 Da, carbamidomethylation, oxidation on methionine, -2H on cysteine (disulfide) as variable modifications. A maximum false discovery rate of 0.05 using decoy database search, an ion score cut off of 20 and a minimum of 2 identified peptides were chosen as identification criteria. For detection of disulfide crosslinks the respective database of crosslinked linearized peptides of rhG-CSF was generated by xComb [21] for trypsin and for chymotrypsin digests, allowing inter- and intramolecular crosslinks and up to three missed cleavage sites, and searched with Mascot as described above.

For SDS-PAGE basic equipment and gels derived from Biorad (Mini-PROTEAN TGX 4-20 % resolving gels). Novex® sharp unstained protein standard by Invitrogen was used as molecular weight marker. Gels were stained with Phast Gel Blue R Coomassie R350 stain from GE Healthcare or Bio-Rad silver staining kit (161-0443), following the manufacturer's protocol.

2.5 Confocal Laser Scanning Microscopy

CLSM images were generated with a Leica TCS SPE DM5500 Q microscope at 630-fold magnification. Images were generated by projection of the single z-level images from the particle layer in the preparation. We used ANS as fluorescent dye at 405 nm excitation and 450 to 470 nm emission wavelength. 1 μl of 1 mM ANS was added to 10 μl rhG-CSF solution (1.4 mg/ml). Molecular ANS/rhG-CSF ratio was 13.4:1. For image processing ImageJ by Wayne Rasband, NIH, was used.

2.6 Kinetics and Statistics

Origin 7.5 by OriginLab Corporation was applied for fitting experimental data to the F-W 2-step model and statistical analysis as shown by Morris et al. 2008 [22]. Significance was checked using a two-sample t-test after verification of normal distribution of the data with the Shapiro-Wilk normality test. Significance level was 0.95.

2.7 Materials

rhG-CSF (Filgrastim) was obtained by Sandoz, Austria. All chemicals and reagents were purchased at Carl Roth GmbH + Co. KG, Germany except: glutamic acid, 8-anilino-1-naphthalenesulfonic acid, iodoacetamide, dithiothreitol, ethyl acetate, toluene (Sigma-Aldrich), potassium permanganate, potassium carbonate (Merck).

Chapter 3

Results

3.1 Protein Loss

The exposure of rhG-CSF to elevated temperature over a period of seven days led to no decrease of soluble protein. In SDS-PAGE we found no additional aggregate bands in comparison to the unstressed control (Figure 3.1). Intrinsic fluorescence spectra remained unchanged and no dimer or aggregate peaks were identified in SEC. Even after removal of sorbitol from the bulk solution with dialysis we found the same negative results. (Data not shown)

In agitated samples protein concentration decreased linearly with 21.8 ± 3.8 μg ml^{-1} h$^{-}1$ (Figure 3.2). Solutions became turbid already within two hours of stirring. Loading 30 μg precipitated protein on SDS-PAGE delivered discrete bands over the whole gel with a distance to each other resembling the molecular weight of rhG-CSF. Reduction with 10 mM DTT only partially led to a disintegration of the aggregate bands. Low molecular weight bands remained visible which left the actual involvement of disulfide bridges unclear up to that moment (Figure 3.1).

Bubble aerated samples displayed faster loss of soluble rhG-CSF compared to agitated ones (Figure 3.2). Concentration decreased with a rate of 42.1 ± 11.3 μg ml^{-1} h^{-1}. However, aerated protein solutions only became slightly turbid after six hours of stressing. Turbidity never reached the high levels obtained with stirring. SDS-PAGE delivered the same results as for stirred samples (Figure 3.1). However, the amount of precipitated protein to be loaded on the gel was lower. Due to foam formation a considerable part of precipitate remained attached to the bioreactor wall and was therefore not found in solution. Rinsing of the reactor wall was not performed.

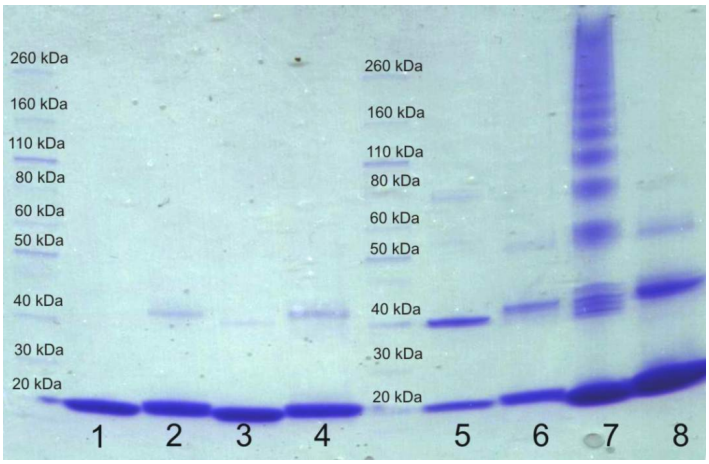


Figure 3.1: SDS-PAGE of stressed samples. Lane 1+2: unstressed rhG-CSF, Lane 3+4: 7 days at 37°C, Lane 5+6: 12 hours bubble aeration, Lane 7+8: 47 hours stirring; all samples in even lanes were reduced with DTT

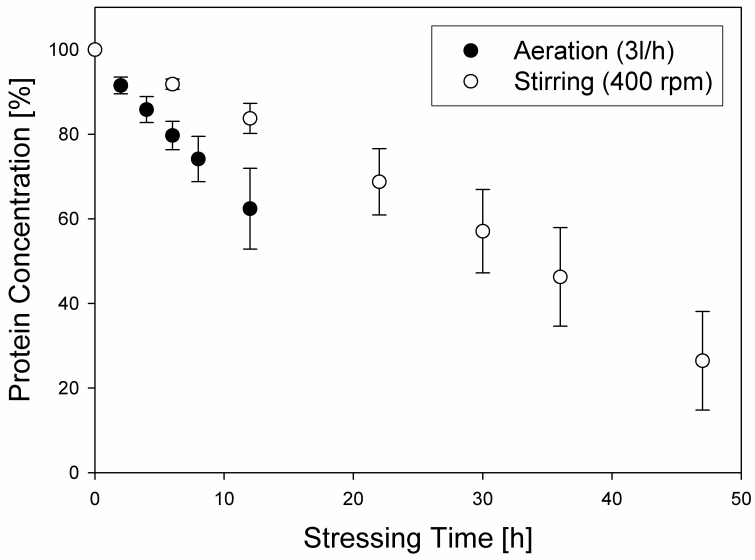


Figure 3.2: Loss of soluble rhG-CSF over stressing time (Mean ± SD). Agitation: -21.8 ± 3.8 μg ml^{-1} h^{-1}; Aeration: -42.1 ± 11.3 μg ml^{-1} h^{-1}; the rate of protein loss was twice as high in aerated samples as in stirred samples

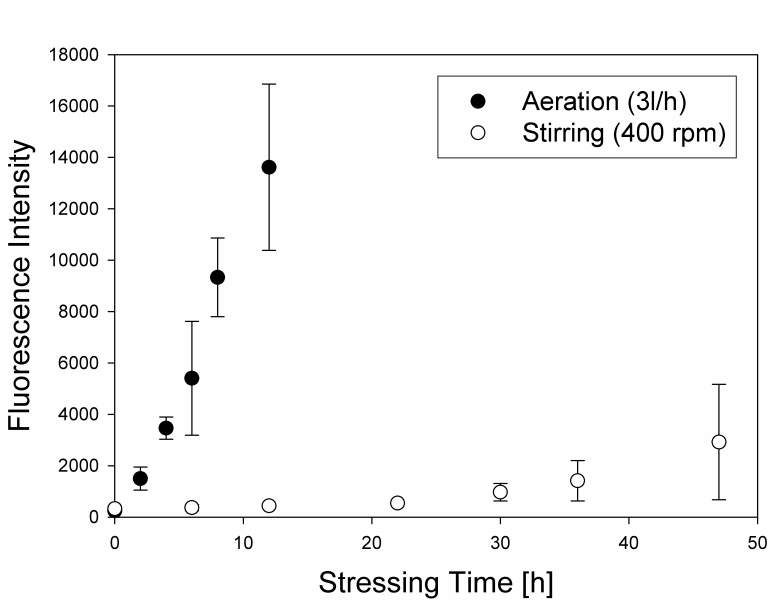


Figure 3.3: ANS fluorescence intensity (arbitrary units) over stressing time (Mean $\pm$ SD). Rapid increase with aeration denotes sizeable structural disruption in rhG-CSF

3.2 Conformational distortion probed by fluorescence spectroscopy

With both stressing conditions, aeration and agitation, we found an increase in ANS fluorescence intensity over time (Figure 3.3). Agitated samples displayed a significant increase after twelve hours of stressing. For aerated samples this was the case already after 2 hours. Throughout the duration of the experiments, rhG-CSF showed significantly elevated ANS fluorescence in aeration experiments compared to stirred ones.

In intrinsic fluorescence assays at 280 nm excitation wavelength we also detected a significant increase in intensity for aerated samples from 6 hours of stressing on (Figure 3.4). No increase within 47 hours was found for agitated samples. Excitation at 295 nm led to the same results except for the overall lower fluorescence intensity. Tryptophane is thus mainly responsible for those intrinsic fluorescence effects. No wavelength shift of emission maximum was observed.

3.3 Determination of soluble aggregate fractions

With SEC we were able to quantify the amount of soluble aggregates compared to that of monomer in solution by determination of the respective peak areas (Figure 3.5). Aggregate

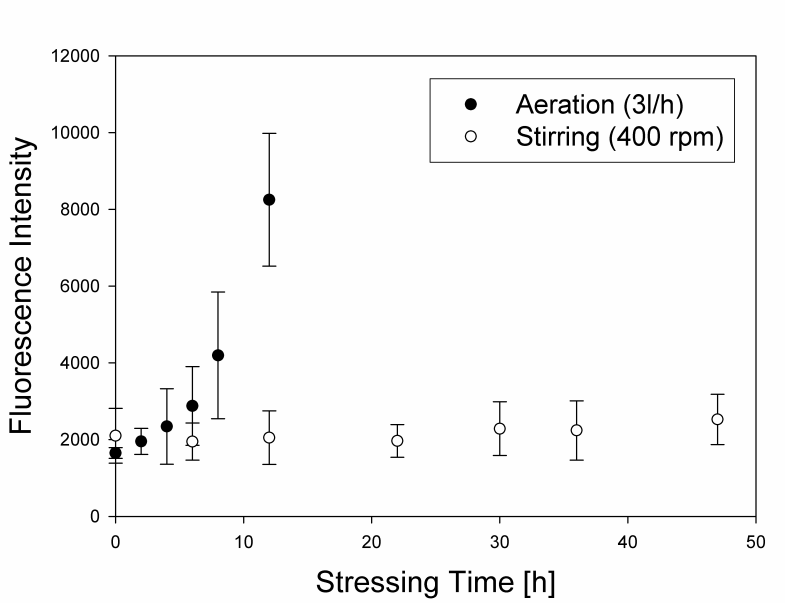


Figure 3.4: Intrinsic fluorescence intensity (arbitrary units) over stressing time (Mean ± SD). A significant increase was only observed with aeration, most likely due to loss of structural flexibility or loss of tryptophane quenching effects

content immediately increased exponentially in aerated samples, whereas in agitated samples a significant rise could only be detected after 22 hours (Figure 3.6). Aerated samples displayed higher contents of soluble aggregates than stirred ones at any time of the experiments. Dimer peaks were only observed in highly aeration-stressed samples. For kinetic analysis dimer fractions were included in aggregate fractions.

3.4 Involvement of disulfide bridges

MS revealed massive rearrangement of disulfide bridges during the aggregation process of rhG-CSF. Table 3.1 features all disulfide bonds that were discovered in MS. Throughout the aggregate bands in SDS-PAGE we found the native disulfide bridges C^{37}-C^{43} and C^{65}-C^{75} as well as the free cysteine C^{18}. We also found free forms of natively bound cysteines in all stirred samples (Table 3.2). In dimer and pentamer bands we found disulfide bridges between C^{18} and C^{75}. Aerated samples also exhibited newly formed disulfide bridges in addition to the native ones. We found C^{18}-C^{18}, C^{18}-C^{37} and C^{18}-C^{75} bonds from dimer to octamer bands. In addition to C^{18} we only found C^{43} to be reduced in aerated samples. All mentioned fragments exhibited a mascot ion score above 15, based on a random probability of 0.05 or below. False discovery rate was in all cases 0.00. The MS spectra of four samples can be found in Supplementary Information A-D.

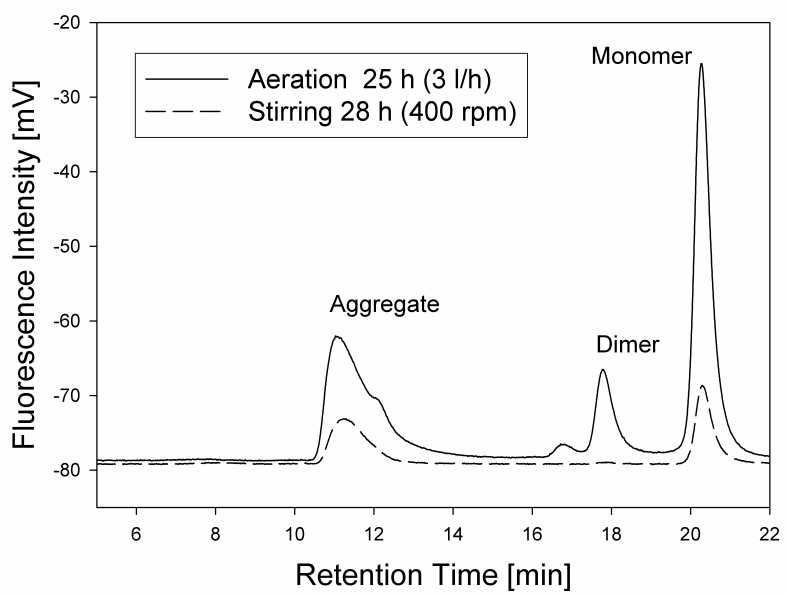


Figure 3.5: SEC-chromatogram examples from aeration and stirring experiments. The area under aggregate peaks was compared with monomer peaks to determine aggregate fraction. Dimer peaks only appeared in very highly aeration-stressed samples

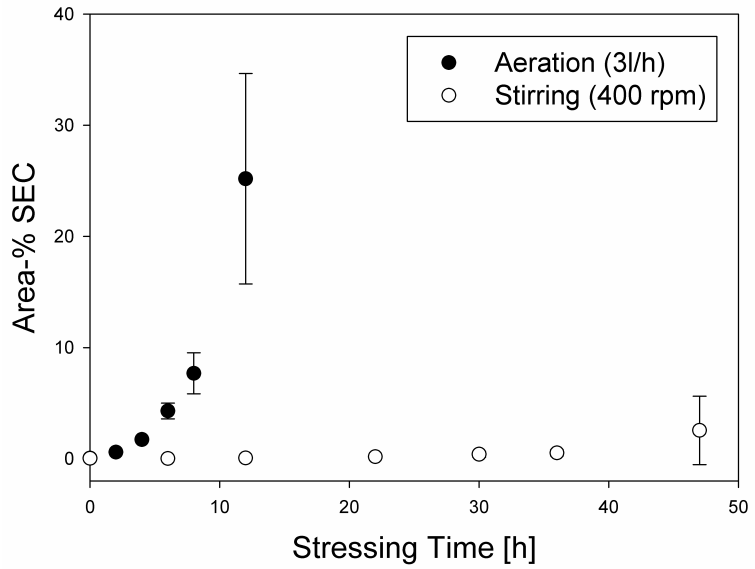


Figure 3.6: Fractions of soluble aggregates found with SEC in stirred and bubble aerated samples (Mean $\pm$ SD). Dimer peaks are not included

Table 3.1: Native and novel disulfide bonds detected with MS after stirring and aeration. The characters in brackets indicate the level of oligomerization; For bold entries MS spectra are provided in the supplementary information

Sequence	Ion score	Mass error [Da]	SS bridge position	Aeration Monomer	Aeration Dimer	Aeration Trimer	Aeration Tetramer	Aeration Polymer	Stirring Monomer	Stirring Dimer	Stirring Trimer	Stirring Tetramer	Stirring Polymer
CATY KLCHPEEL [a] [c]	39	0.0046	C^{37}-C^{43}	x	x	x	x	x (5, 6, 7, 10)			x	x	x (5)
CATYKL CHPEELVL [a] [a]	39	0.0032	C^{37}-C^{43}	x	x	x		x (6)					
CATYKL CHPEELVLL [a] [c]	53	0.0043	C^{37}-C^{43}	x	x	x		x (6)					
KLCHPEELVL CATY [a] [c]	20	0.0053	C^{37}-C^{43}	x	x	x							
QLAGCL SSCPSQAL [a] [c]	24	0.0011	C^{65}-C^{75}	x	x	x	x	x (5, 6, 7, 8, 9, 10)	x	x	x	x	x (5, 6)
SSCPSQAL QLAGCLSQL [a] [c]	21	0.0067	C^{65}-C^{75}	x	x	x	x	x (6,9)	x	x	x	x	x (5)
SSCPSQALQL AGCLSQLHSGL [a] [c]	43	0.0115	C^{65}-C^{75}	x	x								
APLSSCPSQAL QLAGCL [a] [c]	23	0.0098	C^{65}-C^{75}			x							
APLSSCPSQAL QLAGCLSQL [a] [c]	16	0.0084	C^{65}-C^{75}		x								
CLEQVR CLEQVR [b] [d]	52	0.0024	C^{18}-C^{18}		x			x (6, 7)					
LLKCL EQVRKIQGDGAALQEKLCATY [b] [c]	18	0.129	C^{18}-C^{37}					x (6)					
LLKCL QLAGCL [b] [c]	20	0.0026	C^{18}-C^{75}			x	x	x (5, 6, 7, 8)		x			x (5)
LLKCL QLAGCLSQL [b] [c]	20	0.0046	C^{18}-C^{75}					x					

[a] Native disulfide bond

[b] Novel disulfide bond due to stress

[c] Digested with Chymotrypsin

[d] Digested with Trypsin

Table 3.2: Reduced cysteine residues detected with MS after stirring and aeration. Cysteine residue without carbamidomethyl tag represent artefacts from MS analysis or sample preparation. The characters in brackets indicate the level of oligomerization

Sequence	Ion score	Mass error [Da]	Position	Aeration Monomer	Aeration Dimer	Aeration Trimer	Aeration Tetramer	Aeration Poly-mer	Stirring Monomer	Stirring Dimer	Stirring Trimer	Stirring Tetramer	Stirring Poly-mer
KIQGDGAALQEKCATYK [a d]	92	0.0009	C^{37}						x		x		x (6,7)
IQGDGAALQEKCATYK [a d]	106	0.0007	C^{37}								x		x (6,7)
LCATYK [a d]	23	0.0000	C^{37}									x	x (6)
KLCHPEEL [a d]	40	0.0009	C^{43}	x	x			x(9,10)		x	x	x	x(5)
KLCHPEELVL [a c]	58	0.0022	C^{43}		x	x	x	x (5,6,7,8,9,10)		x	x	x	x(5)
SSCPSQALQL [a c]	15	0.0008	C^{65}								x		
AGCLSQL [a c]	40	0.0000	C^{75}							x	x		
QLAGCLSQL [a c]	44	-0.0044	C^{75}								x		
CLEQVRK [b d]	50	0.0007	C^{18}	x					x				
CLEQVR [b d]	30	0.0014	C^{18}	x									
EQVRKIQGDGAALQEKLCATY [b c]	34	0.0021	C^{37}	x									
KLCHPEEL [b c]	37	0.0011	C^{43}	x	x			x(10)	x	x			
KLCHPEELVL [b c]	45	0.0028	C^{43}		x								
SSCPSQALQL [b c]	16	0.0001	C^{65}	x									

[a] Reduced due to stressing - carbamidomethyl tag

[b] Reduced during analysis or sample preparation - no carbamidomethyl tag

[c] Digested with Chymotrypsin

[d] Digested with Trypsin

3.5 Size and shape of insoluble aggregates

Insoluble ANS stained aggregates could be depicted successfully with confocal laser scanning microscopy. Particles did not become visible earlier than after four hours of stressing. We therefore analyzed images from samples taken after four, six and twelve hours of stressing. Three to five images were taken from randomly chosen spots of the preparation. The total particle areas and particle counts were all normalized to a total observation area of 1 mm^2. For both stressing conditions total area and total particle count increased with stressing time.

Due to the high skewness of the particle size distribution towards small particles a comparison of arithmetic means of particle areas would be inappropriate. We therefore present the particle size distributions, as well as the shape distribution in the analyzed samples graphically (Figure 3.7). Looking at the frame between 0 and 1 μm^2 we found between 92.3 and 97.6 % of the total number aggregate particles. While the particle number clearly increased over time, size distribution did not change dramatically. After four hours of aeration most particles had an area between 0.1 and 0.2 μm^2. After that, the maximum shifted back to below 0.1 μm^2. Comparing the stressing conditions we can say that there is a tendency towards smaller particles in aerated samples, where 64.4 % of the particles have an area below 0.2 μm^2. In stirred samples only 56.7 % are smaller than 0.2 μm^2. As shape descriptor for rhG-CSF aggregate particles we chose circularity, as defined in Equation 3.1. A value of 1.0 indicates a perfect circle. As the value approaches 0.0 it indicates an increasingly elongated shape. Under both stressing conditions there was an obvious trend towards non-circular particles over time, although the vast majority of particles remained circular. Additionally, we found higher amounts of non-circular particles in aerated samples, compared to stirred ones.

$$Circularity = \frac{4\pi * area}{perimeter^2} \tag{3.1}$$

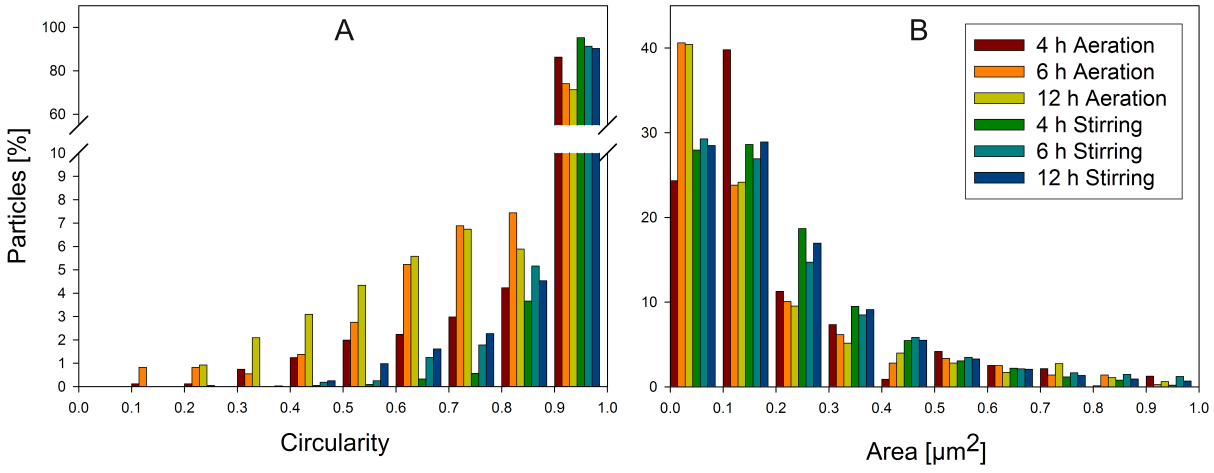


Figure 3.7: A: Distribution of rhG-CSF aggregate particles from stirring and aeration; 1 = perfect circle, 0 = stick-shape. B: Area distribution of rhG-CSF aggregate particles from stirring and aeration

Chapter 4

Discussion

4.1 Aggregation and Precipitation

Incubation at 37°C did not lead to any detectable aggregation of rhG-CSF. As shown before a lower pH value greatly stabilizes rhG-CSF against aggregation [23]. The used pH of 4.4 might therefore obstruct aggregation under these conditions. A reason for this might be the high intrinsic pKa value of the free C^{18}. With H++ it has been calculated to be 9.15 [24]. According to this, at ph 4.4 a considerable higher fraction of free thiols is protonated compared to pH 7. The reduced state is therefore clearly favored at pH 4.4. As we are able to prove the impact of C^{18} on rhG-CSF aggregation in this work, an increased tendency to oxidation is likely to increase stability at lower pH.

The contribution of elevated temperature to aggregation during pharmaceutical processing and storage can thus be considered minimal under optimized pH conditions. Our efforts to induce aggregation by means of stirring and bubble aeration were more successful. With both types of treatment we achieved considerable loss of soluble rhG-CSF, though aeration turned out to be the more effective way.

Quantification of soluble aggregates with SEC confirmed these findings. Aggregate fractions in aerated samples increased exponentially from the beginning, whereas in stirred ones a significant increase could only be detected after 22 hours. Kiese and coworkers reported similar findings for an IgG1 antibody. Shaking, comparable with aeration due to the headspace in shaker flasks, induced more soluble aggregates. Stirring resulted in higher turbidity and insoluble particle amounts [25].

Even after 47 hours, when rhG-CSF concentration reached a level comparable with concentra-

tion in highly aeration stressed samples, we only detected very little amounts of soluble aggregates compared to aeration. In SEC as well as in silver stained SDS-PAGE (data not shown).

It seems very likely that surface effects cause severe structural perturbance in rhG-CSF in a way that hydrophobic protein sections are translocated from the protein's core to the surface. Therefore, exposure of rhG-CSF to air-liquid-interfaces might be a more powerful trigger for aggregation. Stirring on the other hand seems to greatly favor protein precipitation and formation of insoluble aggregates. Fundamentally different pathways might be involved aggregation caused by stirring and aeration.

4.2 Exposure of hydrophobic domains

The hypothesis of different aggregation pathways in stirring and aeration is supported by our results in ANS fluorescence. They indicate a considerable increase of hydrophobic amino acids accessible for ANS, especially in aerated samples. Our results from intrinsic fluorescence correspond well with ANS fluorescence. The increase in fluorescence intensity over aeration time reports conformational modifications. The effect could be explained by loss of structural flexibility or tryptophane quenching effects of neighboring amino acids. The negative intrinsic fluorescence results in stirred samples must therefore demonstrate that only comparatively little conformational changes occur under these conditions. No wavelength shift in emission maximum was observed, since both tryptophane residues are already solvent exposed in native state.

4.3 Disulfide Rearrangement

Mass spectrometry enabled detailed insight into the relevance of disulfide bond breakup and formation in rhG-CSF aggregation. In stirred and in aerated samples we found free cysteine residues from low to high molecular weight aggregates. Not only C^{18}, also C^{43} could be detected unbound throughout the samples. In stirred ones we additionally found free C^{37} and C^{75} thiols. Free C^{65} was only detected in stirred trimer bands. The only new disulfide bridge in stirred samples was C^{18}-C^{75}. Lying on opposite helices there is a distance of 13.9 to 22.5 Å between the two sulfur atoms. Provided sizeable structural disturbance one could consider it as an intramolecular disulfide bond as well as an intermolecular one. Solely from MS results we cannot deduce if those fragments were part of a single protein molecule or not.

The same newly formed disulfide bond could be detected in aerated samples, just like C^{18}-

C^{37} and C^{18}-C^{18}. For C^{18}-C^{37} being an intramolecular disulfide bond is actually very unlikely, as the two residues are located on the same helix in a distance of 27.5 to 31.0 Å. And C^{18}-C^{18} can of course only derive from different molecules. Therefore, we can prove that intermolecular disulfide bridges are at least involved in surface induced rhG-CSF aggregation process. But the fact that we found those newly formed bonds particularly in higher molecular weight aggregate bands and not one of them in monomer bands clearly suggests that also the other disulfide bonds contribute to aggregation.

Furthermore, we can conclude that different types of stressing cause different disulfide bond patterns. Stirring produced more free cysteine residues, whereas more new disulfide bridges were found in aerated samples. Our fluorescence results confirm that aeration causes more conformational perturbance than stirring under the investigated conditions. A certain level of conformational change is certainly necessary for the formation of new disulfide bridges. Thus, it is likely that more structural rearrangement increases the chance to produce them.

4.4 Insoluble Particle Distributions

CLSM images revealed size and shape distributions of insoluble rhG-CSF aggregates. Comparison and evaluation of total particle areas and particle counts cannot be presented. Foam formation during bubble aeration led to substantial attachment of protein on the bioreactor wall. This was not the case in stirred reactors. Nevertheless, a comparison of particle size and shape was possible, though especially large aeration derived particles could have been excluded from the investigation by that.

We did not observe the formation of protein fibrils or particles of comparable elongated shape as they could have been expected from earlier CLSM images and electron microscopy [26][27][28] [29]. Neither in aerated nor in stirred samples. The particle majority was compact and circularly shaped. An amorphous rather than a fibrillary aggregation behavior is most probable. We only found a slight tendency towards elongated shapes in aerated samples, which was more pronounced in six hours and twelve hours-samples.

The distribution of particle size in stirred samples did not change exceptionally over time. For aeration, the maximum of observed particles shifted from between one and two μm^2 to below one μm^2 after four hours. Higher soluble protein concentrations seem to favor the formation of larger aggregate particles, at least in aerated samples.

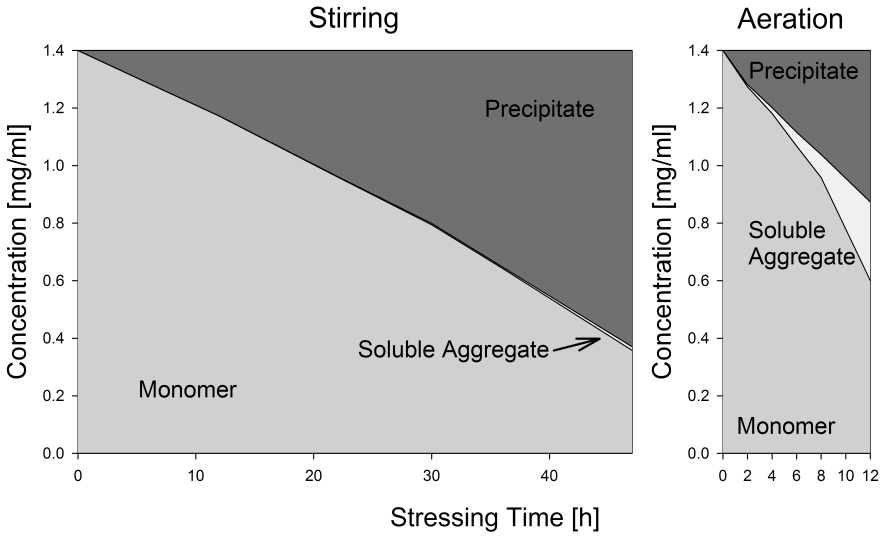


Figure 4.1: Mass fractions obtained during stressing experiments

4.5 Aggregation Kinetics

A kinetic description of soluble protein concentration could easily be accomplished with linear fits ($R^2 > 0.95$). Protein loss occurred twice as fast in aerated samples compared to stirred ones (Figure 3.2). However, the formation of soluble aggregates was not a linear process. Compared to the total amount of protein, soluble aggregates comprised only small fractions (Figure 4.1). But removal of the precipitate and thus looking only at the soluble protein provided an alternate view on kinetics. Normalizing the remaining soluble rhG-CSF to 100 % we consistently observed a lag phase prior to an exponential rise of aggregate fraction in SEC (Figure 3.6).

For the kinetic comparison of these processes we employed the Finke-Watzky mechanism of nucleation followed by autocatalytic surface growth (F-W 2-step, Scheme 4.1) [30]. It proved to be applicable to a wide range of aggregation data from non-prion research literature [22]. It describes aggregation only with two pseudoelementary steps, namely slow continuous nucleation (k1) and typically fast autocatalytic surface growth (k2). Despite this simplification it was possible to fit aggregation data obtained from SEC yielding excellent coefficients of determination ($R^2 \geq 0.995$). We employed the sums of dimer and aggregate fractions for fitting the species *B* and presumed an endpoint of 100 % aggregation (Figure 4.2). As *A* is defined as complementary species to *B*, and as monomer and aggregate fractions are also normalized to 100 %, the fitting of monomer data from SEC to *A* was successful as well and yielded the same rate constants.

We determined k1 = 1.5*10^{-3} h^{-1} and k2 = 4.0*10^{-3} % h^{-1} for aerated samples. Stirring pro-

$$A \xrightarrow{k_1} B$$
$$A + B \xrightarrow{k_2} 2B$$

Scheme 4.1: Finke-Watzky mechanism of nucleation followed by autocatalytic surface growth (F-W 2-step); k1 represents nucleation rate, k2 represents autocatalytic surface growth

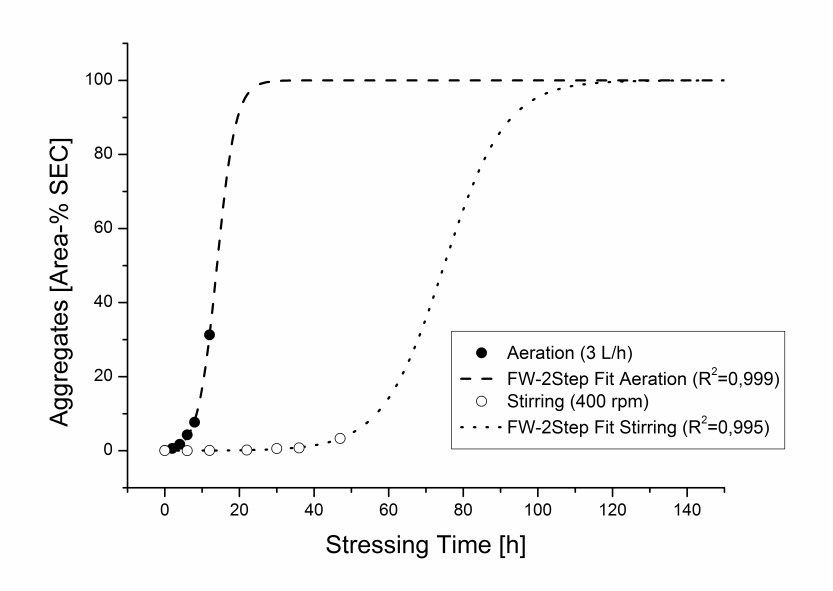


Figure 4.2: Fit of the F-W 2-step mechanism to SEC aggregation data; excellent coefficients of determination ≥ 0.995 were obtained; aggregation endpoint was assumed at 100 %.

vided k1 = 1.0*10^{-5} h^{-1} and k2 = 1.2*10^{-3} % h^{-1}. The putative nucleation rate under the given surface induced stress conditions is thus 150 times higher than under mechanical stirring stress. Aggregate growth rate is only about three times higher. If we assume that nucleation rate represented monomer activation in the applied model, we see that the two types of stressing mainly differ in that first aggregation step. Using the F-W 2-step model we can thus deduce that initiation of rhG-CSF aggregate formation in solution is massively accelerated through bubble aeration.

A striking limitation of the F-W 2-step mechanism is its simplicity. Numerous elementary steps during aggregation are condensed to only two steps leading to a very fuzzy separation of the proposed species [31]. As the nucleus-state is not further characterized in this model we can only state that the initial pseudoelementary step in aggregation, which certainly contains partial protein unfolding, is substantially accelerated under aeration.
Furthermore, this model does not describe or implicate precipitation in any kind. As we observe constant loss of soluble protein from the beginning of our experiments, the obtained nucleation and growth rate cannot be considered accurate. They ougth to be considerably higher to yield such amounts of precipitate. Ignoring precipitation in our kinetic models would presume the existence of entirely different mechanisms for soluble and insoluble aggregate formation. However, we did not succeed in incorporating precipitation into the F-W 2-step mechanism. The introduction of a third rate constant representing precipitation only led to very poor fits to our data. Therefore, we have to emphasize that the employed model only describes formation of soluble aggregates, irrespective of precipitation effects.

Numerous other models try to describe protein aggregation kinetically, also with respect to unfolding, condensation and precipitation [31]. An expanded Lumry-Eyring nucleated polymerization model would combine all of those effects and deliver profound information about native and non-native aggregation mechanisms [32]. Measuring monomer loss and weight-average molecular weight one could distinguish if soluble aggregate growth occurred via chain polymerization, aggregate-aggregate condensation, or a combination of both. Up to that moment we were not able to determine weight distributions in our stressed samples. But this might be accomplished using light scattering techniques.

4.6 Conclusions

In line with the results from fluorescence spectroscopy and MS we can conclude that structural expansion of the native protein is a central point in rhG-CSF aggregation under process

conditions. ANS and intrinsic fluorescence confirmed conformational perturbance, especially through bubble aeration. Massive disulfide bond rearrangement was detected in aeration and stirring, which must be associated with a certain level of structural change. Process kinetic analysis suggested that structural transition is the key factor for rhG-CSF aggregation under the tested conditions, from which bubble aeration turned out to be most effective in both, inducing structural perturbance and triggering aggregate formation. Exposure of rhG-CSF to air-liquid boundaries shall therefore be avoided during storage and production.

Chapter 5

Appendix

In addition to the publication part, the Appendix features detailed information about experimental setup and methodical backgrounds of my research. For SEC we consistently acted in accordance with the SOP provided by Sandoz [33].

5.1 Stirring

The most convenient and reproducible setup we tried for stirring was using a magnetic stir bar in a miniature bioreactor (Figure 5.1). The reactor had an inner diameter of 32 mm. The stir bar had a length of 30 mm and a diameter of 6 mm. Temperature was adjusted to 25°C with a Julabo F25 refrigerated/ heating circulator. Temperature was checked repeatedly with the temperature sensor connected to the IKA RCT Basic magnetic stirrer.

It was not possible to provide fluid dynamic parameters, such as energy input or stirrer power number for that system. The effects of magnetic stir bars have hardly been characterized, most probably due to their limited use in process technology and their unfavorable friction properties. Thus, we can only describe fluid dynamics within our system using the stirring velocity.

A better description would have been possible using a top-mounted magnetic stirrer (Figure 5.2). With that system we observed extreme fluctuation in aggregation velocities at the same stirring speeds. The rotational speed of the IKA RCT Basic magnetic stirrers was checked using a laser tachometer (Voltcraft DT-10L). The reactors and stirrers needed to be cleaned manually. Residues of the cleaning agent (Hellmanex III, Hellma Analytics), which was necessary to remove precipitated protein, or residues of the rhG-CSF itself might have influenced aggre-

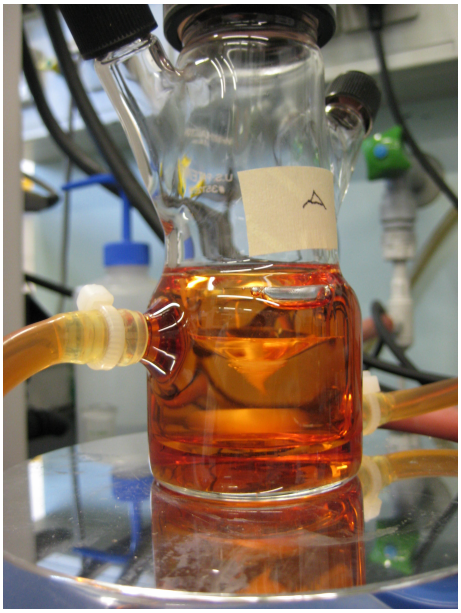


Figure 5.1: A magnetic stir bar in a miniature bioreactor was used for stirring experiments. 15 ml of 1.39 $\pm$ 0.03 mg/ml rhG-CSF solution was stressed for 47 hours at 25°C

gation behavior in an irreproducible way. Nevertheless, even extensive cleaning followed by thorough water rinsing and drying did not improve reproducibility.

Up to that moment we were not able to determine the reason for the poor reproducibility of these experiments. One could think of an uneven stirring behavior of the top mounted stirrers. Actually, we observed small horizontal jiggling movements during stirring. Since we were not able to fix that error, we cannot clarify the significance of our finding. Thus, the only remaining opportunity was to switch to the stir bar system, despite its disadvantages.

5.2 Aeration

We aerated the rhG-CSF solution with compressed air from our in-house system. Pressure there is held around 8*10^5 Pa. To ensure a constant air flow of up to 5 liters per hour a pressure reducer (Stasto R-M14-08-R) was installed. Without that device a massive drop of flow rate over a period of several hours could not be prevented.

In spite of pre-humidification of the air stream, solvent evaporation during bubble aeration could not be prevented completely. We still observed a loss of up to 10 % (v/v) over a period of twelve hours. Longer aeration experiments would inevitably lead to substantial bias in the aggregation results. We considered dosed addition of buffer or water during the experiment.

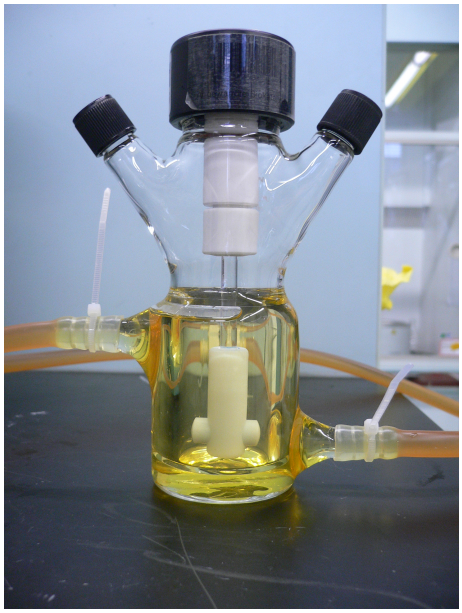


Figure 5.2: Alternative stirring setup using a top-mounted stirrer

But due to unclear interference with the stressing method, such as putative osmotic changes, this was not implemented yet.

5.3 UV Spectrophotometry

Concentration of soluble rhG-CSF was determined spectrophotometrically at 280 nm. A molar extinction coefficient of 15,720 M^{-1} cm^{-1} was obtained from ProtParam tool on the ExPASy Server [34] on the basis of Filgrastim sequence. Molecular weight was 18.8 kDa. We received rhG-CSF from Sandoz in a concentration of 1.9 mg/ml. This was checked using the Bradford method [35]. A calibration of the method was performed in six steps between 1.9 and 0.475 mg/ml. The mean deviation from the expected values accounted for $\pm$ 0.84 %. A volume of 80 μl was analyzed in Plastibrand micro UV-cuvettes. As blank we used 10 mM glutamate buffer, pH 4.4, containing 5 % (w/v) sorbitol. Glutamic acid and sorbitol were dissolved in HPLC-grade water and pH was adjusted with 5 M NaOH. The buffer was filtered through syringe filters with a pore size of 45 μm to reduce light scattering. We are referring to that preparation whenever glutamate buffer is mentioned in that work. Every sample was measured in duplicates.

5.4 Fluorescence Spectroscopy

Fluorescence signals are very sensitive to temperature changes. Therefore, we aimed at keeping the temperature constant at 25°C throughout the flourescence analysis and before. We also scheduled all fluorescence measurements on the day after sampling. This guaranteed constant conditions concerning potential continuing aggregation and aggregate dissociation. rhG-CSF concentration was adjusted to 92 μg/ml by dilution with glutamate buffer. This concentration provided low fluorescence for unstressed samples together with acceptable noise. The appropriate volume of glutamate buffer was placed into a 1.5 ml tube and kept for at least 20 min at 25°C. Then, 20 minutes before measurement, the respective volume of cooled rhG-CSF solution was added in order to allow warming. 500 μl of that solution were placed in an appropriate quartz fluorescence cuvette and employed for measurement.

Excitation occurred at 280 and 295 nm respectively. Excitation and emission slits were set to 5 μm. Emission spectra were recorded between 300 and 500 nm. Photomultiplier voltage was adjusted to 950 V. Three spectra were recorded per sample and smoothed three times to 50 points with the Savitsky-Golay method, smoothing order 3. The average of emission at 340 nm was regarded as result.

Immediately after measurement of intrinsic fluorescence 10 μl of 1 mM ANS were added to obtain a molecular ANS/rhG-CSF ratio of 4:1. After mixing by repeated up-and-down-pipeting and five minutes of incubation the measurement was started.

ANS fluorescence was induced at 388 nm. Emission spectra were recorded from 400 to 500 nm. Other settings and the smoothing method remained unchanged. Results were obtained by the average of the three fluorescence emission values at 470 nm. In case of very highly stressed samples, fluorescence exceeded the maximal detection limit. In that case photomultiplier voltage was reduced to 700 V. Comparison of eleven samples revealed a factor of signal amplification of 7.5 $\pm$ 0.9. 700 V signals were therefore multiplied with 7.5. A dilution of the samples in order to reduce the fluorescence emission signal was considered inappropriate due to dissociation effects of reversible rhG-CSF aggregates, repeatedly observed before [36][23].

For immediate reuse of the quartz cuvette it was flushed with deionized water and 70 % (v/v) ethanol followed by drying in a compressed air stream.

5.5 Size exclusion HPLC

SEC was carried out following the standard operating procedure provided by Sandoz which cannot be presented here [33]. The analysis was performed strictly in the same week as the stressing experiments in order to avoid result corruption during storage. Only a relative quantification of aggregate peak area compared to monomer peak area could be performed due to the lack of reference values for calibration.

5.6 Dialysis and thin layer chromatography (TLC)

For the removal of sorbitol from the Sandoz bulk solution we employed micro-dialysis through a Spectra/Por 4 membrane, MWCO 12-14,000. 15 ml of the bulk solution were placed into the sealed micro-dialysis tube. At 4°C it was put into a stirred beaker glass with sorbitol-free glutamate buffer. Dialysis took place over a period of approximately 40 hours with three to four times repeated exchange of sorbitol-free dialysis buffer.

Preliminary comparison of buffer exchange methods was carried out with recombinant human growth hormone (hGH). Dialysis was compared with NAP-10 sephadex gravity flow columns (GE Healthcare) and Vivaspin 6 ml centrifugal concentrators (MWCO = 10,000, Sartorius). Dialysis showed highest yields of protein concentration compared with the other methods. Only 4 % of hGH were lost during dialysis , compared to 13 % with NAP-columns and 22 % with Vivaspin. We considered the rather gentle conditions during dialysis to be another advantage of the method.

Dialysis performance was checked with TLC for sorbitol. Using glass capillarys 3 μl of the analyzed sample were applied on Merck silica gel F254 aluminum sheets. After drying the mobile phase was allowed to infiltrate the silica sheets for 30 minutes. Mobile phase contained 51 % (v/v) isopropyl alcohol, 26 % (v/v) ethyl acetate, 13 % (v/v) toluene and 10 % (v/v) deionized water. The sheets were dried and sprayed with staining solution, containing 3 g $KMnO_4$, 20 g K_2CO_3, 5 ml 5 % (w/v) NaOH and 300 ml deionized water. Negative staining developed after short heating on a heat plate. Figure 5.3 shows a stained TLC sheet after micro-dialysis.

Figure 5.3: Thin layer chromatography to check dialysis performance; lane 1: 5 % (w/v) sorbitol positive control; lane 2: undialyzed rhG-CSF bulk solution containing 5 % (w/v) sorbitol; lane 3: dialyzed rhG-CSF solution

5.7 SDS-PAGE

To gain impressions of aggregate content in stressed samples and to prepare samples for MS SDS-PAGE was employed. For these purposes we did not remove rhG-CSF precipitates from the solution. By centrifugation a considerable amount of soluble aggregates or aggregates reversible upon addition of SDS are removed with the precipitate. After 1+1 addition of loading dye (20 mM K_2HPO_4, 6 mM EDTA, 6 % (w/v) SDS, 10 % (w/v) Glycerol, 0.05 % (w/v) Bromophenol blue, pH 8.0) and application of 20 μl to the gel an amount of 14 μg stressed rhG-CSF was analyzed per lane.

For preparative SDS-PAGE prior to MS high protein amounts were desired. So unlike the other aeration experiments, bioreactor wall was rinsed with the sample solution to accumulate precipitated rhG-CSF before sampling. After one minute centrifugation of the stressed samples at 1,500 g only the precipitate was employed for SDS-PAGE. The protein amount applied in SDS-PAGE could only be estimated indirectly from the soluble protein concentration. For stirred and aerated samples we reached approximately 30 μg per lane. These amounts might seem quite high, but due to the high content of insoluble aggregates only a small fraction could actually penetrate the gel.

For maximizing the range of visible aggregates we used precast gradient gels (Mini-PROTEAN

TGX 4-20 % Resolving Gels). Tenfold concentrated electrode was prepared with 30.3 g Tris base, 144.0 g Glycine and 10.0 g SDS on one liter deionized water. Gels were run at 180 V for approximately 40 minutes. That is until stain bands had almost run out of the gel. By default we used Phast Gel Blue R Coomassie R350 stain from GE Healthcare. For one hour the gels were left in staining solution, prepared following the manufacturer's instructions. Destaining occurred over night or by repeated exchange of destaining solution until desired level of contrast was reached. Destaining solution contained 30 % (v/v) ethanol and 10 % (v/v) acetic acid in deionized water. Preserve solution for short term storage of the gel consisted of 10 % (v/v) acetic acid and 13 % (v/v) glycerol.

In the case of removal of the precipitate from stressed samples, aggregate bands could only be visualized with silver staining. Best results were obtained using the Bio-Rad Silver Staining Kit (161-0443). Gels were treated as suggested in the 'Modified Silver Stain Protocol' from the manufacturer's guide. Protein amounts loaded on the respective gels ranged from 14 to 4 μg per lane depending on the degree of stressing.

5.8 Mass spectrometry

For reduction of disulfide bonds rhG-CSF precipitate was subjected to 10 mM DTT. After preparative SDS-PAGE we often noticed incomplete disintegration of aggregate bands. Concentration of DTT might be too low to reduce all disulfide bridges. rhG-CSF concentration accounted for approximately 27 μM. With the protein amounts we obtained from 5 ml bubble aerated solution we would at least recommend three times the concentration of DTT for following experiments. The same measure appears reasonable for IAA.

5.9 Confocal Laser Scanning Microscopy

With CLSM we were able to visualize particles down to a diameter of 0.33 μm. A 630fold magnification could be reached with the used Leica TCS SPE DM5500 Q. Stressed samples were employed for microscopy without further purification. 10 μl of rhG-CSF sample were mixed with 1 μl 1 mM ANS solution and incubated for five minutes. Molecular ANS/rhG-CSF ratio was 13.4:1. This solution was applied on a glass slide. To prevent leakage, evaporation or further disturbance the used cover glass was sealed with simple nail polish.

With CLSM images from multiple z-levels can be generated. We took scans from the whole fluid layer in the preparation. In Leica LAS AF software these images were then projected into

a single one, containing the collected information of the stacked images.

With ImageJ software (Wayne Rasband, NIH) we were able to describe quantity, area and shape of the aggregate particles. For that purpose binary images were generated. A black-white-threshold had to be adjusted manually. Compared to actual particle measures, slight discrepancies, especially with small particles, had to be accepted. Distributions could then be generated with ImageJ.

5.10 Materials

If not indicated otherwise, all chemicals derived from Roth GmbH + Co. KG, Germany except: ethyl acetate (Sigma-Aldrich), EDTA (Serva). hGH was provided by Sandoz

List of Figures

1.1 NMR structure of human G-CSF 2

3.1 SDS-PAGE of stressed samples 10
3.2 Loss of soluble rhG-CSF over stressing time 10
3.3 ANS fluorescence intensity over stressing time 11
3.4 Intrinsic fluorescence intensity over stressing time 12
3.5 SEC-chromatogram examples from aeration and stirring experiments 13
3.6 Fractions of soluble aggregates found with SEC in stirred and bubble aerated samples 13
3.7 Distribution of rhG-CSF aggregate particles from stirring and aeration 17

4.1 Mass fractions obtained during stressing experiments 21
4.2 Fit of the F-W 2-step mechanism to SEC aggregation data 22

5.1 A magnetic stir bar in a miniature bioreactor was used for stirring experiments . 26
5.2 Alternative stirring setup using a top-mounted stirrer 27
5.3 Thin layer chromatography to check dialysis performance 30

Bibliography

[1] C. Ross and M. Poirier, "Protein aggregation and neurodegenerative disease.," *Nature medicine*, vol. 10, pp. S10–S17, 2004. 1

[2] I. Weiss, F. William, T. Young, and C. Roberts, "Principles, approaches, and challenges for predicting protein aggregation rates and shelf life," *Journal of pharmaceutical sciences*, vol. 98, no. 4, pp. 1246–1277, 2009. 1

[3] H. Mahler, W. Friess, U. Grauschopf, and S. Kiese, "Protein aggregation: Pathways, induction factors and analysis," *Journal of pharmaceutical sciences*, vol. 98, no. 9, pp. 2909–2934, 2009. 1

[4] M. Manning, D. Chou, B. Murphy, R. Payne, and D. Katayama, "Stability of Protein Pharmaceuticals: An Update," *Pharmaceutical research*, vol. 27, no. 4, pp. 544–575, 2010. 1

[5] B. Chang and S. Hershenson, *Practical approaches to protein formulation development.*, vol. 13. Kluwer Academic/Plenum Publishers, 2002. 1

[6] A. Rosenberg, "Effects of protein aggregates: an immunologic perspective," *The AAPS journal*, vol. 8, no. 3, pp. 501–507, 2006. 1

[7] A. Fernandez-Escamilla, F. Rousseau, J. Schymkowitz, and L. Serrano, "Prediction of sequence-dependent and mutational effects on the aggregation of peptides and proteins," *Nature biotechnology*, vol. 22, no. 10, pp. 1302–1306, 2004. 1

[8] G. Tartaglia and M. Vendruscolo, "The Zyggregator method for predicting protein aggregation propensities," *Chemical Society Reviews*, vol. 37, no. 7, pp. 1395–1401, 2008. 1

[9] C. Hill, T. Osslund, and D. Eisenberg, "The structure of granulocyte-colony-stimulating factor and its relationship to other growth factors," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 90, no. 11, p. 5167, 1993. 2

[10] K. Welte, J. Gabrilove, M. Bronchud, E. Platzer, and G. Morstyn, "Filgrastim (r-metHuG-CSF): the first 10 years," *Blood*, vol. 88, no. 6, p. 1907, 1996. 2, 3

[11] T. Zink, A. Ross, K. Lueers, C. Cieslar, R. Rudolph, and T. Holak, "Structure and dynamics of the human granulocyte colony-stimulating factor determined by NMR spectroscopy. Loop mobility in a four-helix-bundle protein," *Biochemistry*, vol. 33, no. 28, pp. 8453–8463, 1994. 2

[12] S. Raso, J. Abel, J. Barnes, K. Maloney, G. Pipes, M. Treuheit, J. King, and D. Brems, "Aggregation of granulocyte-colony stimulating factor in vitro involves a conformationally altered monomeric state," *Protein Science*, vol. 14, no. 9, pp. 2246–2257, 2005. 3

[13] R. Bartkowski, R. Kitchel, N. Peckham, and L. Margulis, "Aggregation of recombinant bovine granulocyte colony stimulating factor in solution," *Journal of protein chemistry*, vol. 21, no. 3, pp. 137–143, 2002. 3

[14] M. Oh-eda, M. Hasegawa, K. Hattori, H. Kuboniwa, T. Kojima, T. Orita, K. Tomonou, T. Yamazaki, and N. Ochi, "O-linked sugar chain of human granulocyte colony-stimulating factor protects it against polymerization and denaturation allowing it to retain its biological activity.," *Journal of Biological Chemistry*, vol. 265, no. 20, p. 11432, 1990. 3

[15] M. Hasegawa, "A thermodynamic model for denaturation of granulocyte colony-stimulating factor: O-linked sugar chain suppresses not the triggering deprotonation but the succeeding denaturation," *Biochimica et Biophysica Acta (BBA)-Protein Structure and Molecular Enzymology*, vol. 1203, no. 2, pp. 295–297, 1993. 3

[16] V. Gervais, A. Zertal, and H. Oschkinat, "NMR Investigations of the Role of the Sugar Moiety in Glycosylated Recombinant Human Granulocyte-Colony-Stimulating Factor," *European Journal of Biochemistry*, vol. 247, no. 1, pp. 386–395, 1997. 3

[17] E. Chi, S. Krishnan, B. Kendrick, B. Chang, J. Carpenter, and T. Randolph, "Roles of conformational stability and colloidal stability in the aggregation of recombinant human granulocyte colony-stimulating factor," *Protein Science: A Publication of the Protein Society*, vol. 12, no. 5, p. 903, 2003. 3

[18] Y. Maa and C. Hsu, "Protein denaturation by combined effect of shear and air-liquid interface," *Biotechnology and bioengineering*, vol. 54, no. 6, pp. 503–512, 1997. 3

[19] A. Hawe, M. Sutter, and W. Jiskoot, "Extrinsic fluorescent dyes as tools for protein characterization," *Pharmaceutical research*, vol. 25, no. 7, pp. 1487–1499, 2008. 4

[20] A. Shevchenko, M. Wilm, O. Vorm, and M. Mann, "Mass spectrometric sequencing of proteins from silver-stained polyacrylamide gels," *Anal. Chem*, vol. 68, no. 5, pp. 850–858, 1996. 6

[21] A. Panchaud, P. Singh, S. Shaffer, and D. Goodlett, "xComb: A Cross-Linked Peptide Database Approach to Protein- Protein Interaction Analysis," *Journal of Proteome Research*, vol. 9, no. 5, pp. 2508–2515, 2010. 7

[22] A. Morris, M. Watzky, J. Agar, and R. Finke, "Fitting Neurological Protein Aggregation Kinetic Data via a 2-Step, Minimal/'Ockham's Razor' Model: The Finke- Watzky Mechanism of Nucleation Followed by Autocatalytic Surface Growth," *Biochemistry*, vol. 47, no. 8, pp. 2413–2427, 2008. 8, 21

[23] R. Thirumangalathu, S. Krishnan, D. Brems, T. Randolph, and J. Carpenter, "Effects of pH, temperature, and sucrose on benzyl alcohol-induced aggregation of recombinant human granulocyte colony stimulating factor," *Journal of pharmaceutical sciences*, vol. 95, no. 7, pp. 1480–1497, 2006. 18, 28

[24] J. Gordon, J. Myers, T. Folta, V. Shoja, L. Heath, and A. Onufriev, "H++: a server for estimating pKas and adding missing hydrogens to macromolecules," *Nucleic acids research*, vol. 33, no. suppl 2, p. W368, 2005. 18

[25] S. Kiese, A. Papppenberger, W. Friess, and H. Mahler, "Shaken, not stirred: Mechanical stress testing of an IgG1 antibody," *Journal of pharmaceutical sciences*, vol. 97, no. 10, pp. 4347–4366, 2008. 18

[26] T. Ban, K. Morigaki, H. Yagi, T. Kawasaki, A. Kobayashi, S. Yuba, H. Naiki, and Y. Goto, "Real-time and single fibril observation of the formation of amyloid β spherulitic structures," *Journal of Biological Chemistry*, vol. 281, no. 44, pp. 33677–33683, 2006. 20

[27] D. Bateman, J. McLaurin, and A. Chakrabartty, "Requirement of aggregation propensity of Alzheimer amyloid peptides for neuronal cell surface binding," *BMC neuroscience*, vol. 8, no. 1, p. 29, 2007. 20

[28] T. Dickson and J. Vickers, "The morphological phenotype of [beta]-amyloid plaques and associated neuritic changes in Alzheimer's disease," *Neuroscience*, vol. 105, no. 1, pp. 99–107, 2001. 20

[29] P. Taboada, S. Barbosa, E. Castro, and V. Mosquera, "Amyloid fibril formation and other aggregate species formed by human serum albumin association," *J. Phys. Chem. B*, vol. 110, no. 42, pp. 20733–20736, 2006. 20

[30] M. Watzky and R. Finke, "Transition metal nanocluster formation kinetic and mechanistic studies. A new mechanism when hydrogen is the reductant: slow, continuous nucleation and fast autocatalytic surface growth," *J. Am. Chem. Soc*, vol. 119, no. 43, pp. 10382–10400, 1997. 21

[31] A. Morris, M. Watzky, and R. Finke, "Protein aggregation kinetics, mechanism, and curve-fitting: a review of the literature," *Biochimica et Biophysica Acta (BBA)-Proteins & Proteomics*, vol. 1794, no. 3, pp. 375–397, 2009. 23

[32] C. Roberts, "Non-native protein aggregation kinetics," *Biotechnology and bioengineering*, vol. 98, no. 5, pp. 927–938, 2007. 23

[33] A. Thaler, H. Lerch, U. Krimm, and H. Schabasser, *Determination of purity and identity of EP2006 by size exclusion chromatography*. Sandoz, 5.0 ed., 2009. 25, 29

[34] E. Gasteiger, C. Hoogland, A. Gattiker, S. Duvaud, M. Wilkins, R. Appel, and A. Bairoch, "Protein identification and analysis tools on the ExPASy server," *The proteomics protocols handbook*, pp. 571–607, 2005. 27

[35] M. Bradford, "A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding," *Analytical biochemistry*, vol. 72, no. 1-2, pp. 248–254, 1976. 27

[36] S. Krishnan, E. Chi, J. Webb, B. Chang, D. Shan, M. Goldenberg, M. Manning, T. Randolph, and J. Carpenter, "Aggregation of Granulocyte Colony Stimulating Factor under Physiological Conditions: Characterization and Thermodynamic Inhibition," *Biochemistry*, vol. 41, no. 20, pp. 6422–6431, 2002. 28

Supplementary Information

Supplementary Information A - MS spectrum CATYKL-CHPEELVLL

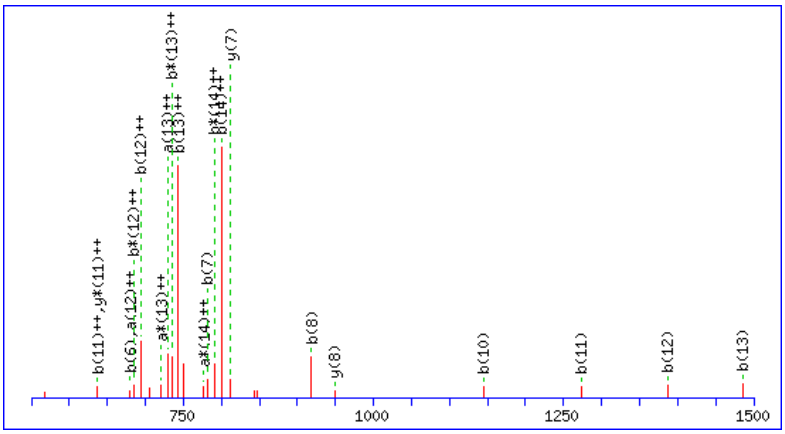
Mascot Search Results

Peptide View

MS/MS Fragmentation of **CATYKLCHPEELVLL**
Found in **P09919_ab_25_7**, a=CATYKL (37-42) cx b=CHPEELVLL (43-51) Granulocyte colony-stimulating factor OS=Homo sapiens GN=CSF3 PE=1 SV=1

Match to Query 401: 1728.862148 from(865.438350,2+) intensity(30208668.0000)
Title: 335: Sum of 9 scans in range 1540 (rt=30.245) to 1643 (rt=31.7574) [E:\Data\RCPE1\RCPE1_II_20100917\Chymotr_01.RAW]
Data file E:\Data\RCPE1\RCPE1_II_20100917\Chymotr_01.RAW

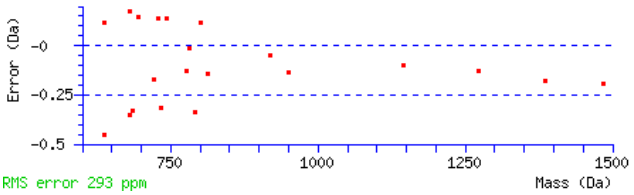
Click mouse within plot area to zoom in by factor of two about that point
Or, Plot from 550 to 1500 Da Full range

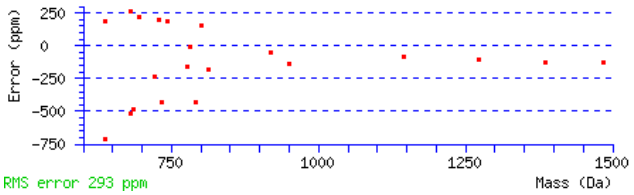


Monoisotopic mass of neutral peptide Mr(calc): 1728.8579

Variable modifications:
C7 : Disulfide bridge (C)
Ions Score: 53 **Expect:** 4.6e-007
Matches (Bold Red): 21/132 fragment ions using 24 most intense peaks

#	a	a^{++}	a*	a^{*++}	b	b^{++}	b*	b^{*++}	Seq.	y	y^{++}	y*	y^{*++}	#
1	76.0215	38.5144			104.0165	52.5119			C					15
2	147.0587	74.0330			175.0536	88.0304			A	1626.8560	813.9316	1609.8295	805.4184	14
3	248.1063	124.5568			276.1013	138.5543			T	1555.8189	778.4131	1538.7923	769.8998	13
4	411.1697	206.0885			439.1646	220.0859			Y	1454.7712	727.8892	1437.7447	719.3760	12
5	539.2646	270.1360	522.2381	261.6227	567.2595	284.1334	550.2330	275.6201	K	1291.7079	646.3576	1274.6813	**637.8443**	11
6	652.3487	326.6780	635.3221	318.1647	**680.3436**	340.6754	663.3171	332.1622	L	1163.6129	582.3101			10
7	753.3422	377.1748	736.3157	368.6615	**781.3372**	391.1722	764.3106	382.6589	C	1050.5289	525.7681			9
8	890.4012	445.7042	873.3746	437.1909	**918.3961**	459.7017	901.3695	451.1884	H	**949.5353**	475.2713			8
9	987.4539	494.2306	970.4274	485.7173	1015.4488	508.2281	998.4223	499.7148	P	**812.4764**	406.7418			7
10	1116.4965	558.7519	1099.4700	550.2386	**1144.4914**	572.7494	1127.4649	564.2361	E	715.4236	358.2155			6
11	1245.5391	623.2732	1228.5126	614.7599	**1273.5340**	**637.2706**	1256.5075	628.7574	E	586.3810	293.6942			5
12	1358.6232	**679.8152**	1341.5966	671.3019	**1386.6181**	**693.8127**	1369.5915	**685.2994**	L	457.3384	229.1729			4
13	1457.6916	**729.3494**	1440.6650	**720.8362**	**1485.6865**	**743.3469**	1468.6599	**734.8336**	V	344.2544	172.6308			3
14	1570.7756	785.8915	1553.7491	**777.3782**	1598.7706	**799.8889**	1581.7440	**791.3756**	L	245.1860	123.0966			2
15									L	132.1019	66.5546			1





NCBI **BLAST** search of CATYKLCHPEELVLL
(Parameters: blastp, nr protein database, expect=20000, no filter, PAM30)
Other BLAST web gateways

All matches to this query

Score	Mr(calc):	Delta	Sequence
53.3	1728.8579	0.0043	CATYKLCHPEELVLL
49.9	1728.8579	0.0043	CATYKLCHPEELVLL

Mascot: http://www.matrixscience.com/

Supplementary Information B - MS spectrum CLEQVR-CLEQVR

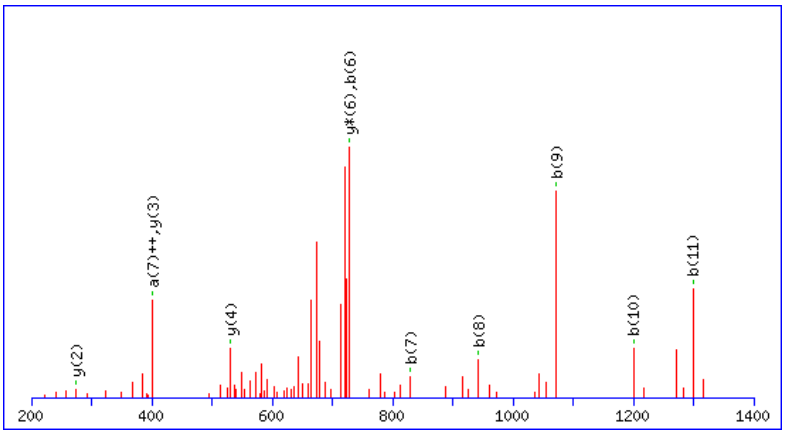
Mascot Search Results

Peptide View

MS/MS Fragmentation of **CLEQVRCLEQVR**
Found in **P09919_aa_3_0**, a=CLEQVR (18-23) cx a=CLEQVR (18-23) Granulocyte colony-stimulating factor OS=Homo sapiens GN=CSF3 PE=1 SV=1

Match to Query 343: 1472.725228 from(737.369890,2+) intensity(23721.1970)
Title: 149: Sum of 2 scans in range 883 (rt=21.692) to 890 (rt=21.8054) [E:\Data\RCPE1\RCPE1_II_20100917\Trypsin_17.RAW]
Data file E:\Data\RCPE1\RCPE1_II_20100917\Trypsin_17.RAW

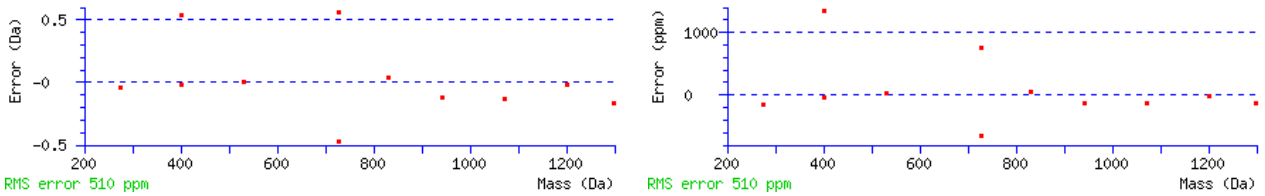
Click mouse within plot area to zoom in by factor of two about that point
Or, Plot from 200 to 1400 Da Full range



Monoisotopic mass of neutral peptide Mr(calc): 1472.7228

Variable modifications:
C7 : Disulfide bridge (C)
Ions Score: 52 **Expect:** 2.9e-007
Matches (Bold Red): 11/120 fragment ions using 18 most intense peaks

#	a	a^{++}	a*	a^{*++}	b	b^{++}	b*	b^{*++}	Seq.	y	y^{++}	y*	y^{*++}	#
1	76.0215	38.5144			104.0165	52.5119			C					12
2	189.1056	95.0564			217.1005	109.0539			L	1370.7209	685.8641	1353.6944	677.3508	11
3	318.1482	159.5777			346.1431	173.5752			E	1257.6368	629.3221	1240.6103	620.8088	10
4	446.2068	223.6070	429.1802	215.0938	474.2017	237.6045	457.1751	229.0912	Q	1128.5942	564.8008	1111.5677	556.2875	9
5	545.2752	273.1412	528.2486	264.6280	573.2701	287.1387	556.2436	278.6254	V	1000.5357	500.7715	983.5091	492.2582	8
6	701.3763	351.1918	684.3498	342.6785	**729.3712**	365.1892	712.3447	356.6760	R	901.4673	451.2373	884.4407	442.7240	7
7	802.3699	**401.6886**	785.3433	393.1753	**830.3648**	415.6860	813.3382	407.1727	C	745.3661	373.1867	**728.3396**	364.6734	6
8	915.4539	458.2306	898.4274	449.7173	**943.4488**	472.2281	926.4223	463.7148	L	644.3726	322.6899	627.3461	314.1767	5
9	1044.4965	522.7519	1027.4700	514.2386	**1072.4914**	536.7493	1055.4649	528.2361	E	**531.2885**	266.1479	514.2620	257.6346	4
10	1172.5551	586.7812	1155.5285	578.2679	**1200.5500**	600.7786	1183.5235	592.2654	Q	**402.2459**	201.6266	385.2194	193.1133	3
11	1271.6235	636.3154	1254.5970	627.8021	**1299.6184**	650.3128	1282.5919	641.7996	V	**274.1874**	137.5973	257.1608	129.0840	2
12									R	175.1190	88.0631	158.0924	79.5498	1



NCBI **BLAST** search of CLEQVRCLEQVR
(Parameters: blastp, nr protein database, expect=20000, no filter, PAM30)
Other BLAST web gateways

All matches to this query

Score	Mr(calc):	Delta	Sequence
52.4	1472.7228	0.0024	CLEQVRCLEQVR
41.1	1472.7228	0.0024	CLEQVRCLEQVR

Mascot: http://www.matrixscience.com/

Supplementary Information C -

MS spectrum SSCPSQALQL-AGCLSQLHSGL

Mascot Search Results

Peptide View

MS/MS Fragmentation of **SSCPSQALQLAGCLSQLHSGL**
Found in **P09919_ab_49_7**, a=SSCPSQALQL (63-72) cx b=AGCLSQLHSGL (73-83) Granulocyte colony-stimulating factor OS=Homo sapiens GN=CSF3 PE=1 SV=1

Match to Query 461: 2097.009848 from(1049.512200,2+) intensity(575969.3200)
Title: 556: Sum of 4 scans in range 3060 (rt=55.2769) to 3096 (rt=55.8466) [E:\Data\RCPE1\RCPE1_II_20100917\Chymotr_02.RAW]
Data file E:\Data\RCPE1\RCPE1_II_20100917\Chymotr_02.RAW

Click mouse within plot area to zoom in by factor of two about that point
Or, Plot from 300 to 1800 Da Full range

Monoisotopic mass of neutral peptide Mr(calc): 2096.9983

Variable modifications:
C3 : Disulfide bridge (C)
Ions Score: 43 **Expect:** 5e-006
Matches (Bold Red): 16/210 fragment ions using 17 most intense peaks

#	a	a^{++}	a*	a^{*++}	b	b^{++}	b*	b^{*++}	Seq.	y	y^{++}	y*	y^{*++}	#
1	60.0444	30.5258			88.0393	44.5233			S					21
2	147.0764	74.0418			175.0713	88.0393			S	2010.9736	1005.9904	1993.9470	997.4771	20
3	248.0700	124.5386			276.0649	138.5361			C	1923.9415	**962.4744**	1906.9150	953.9611	19
4	345.1227	173.0650			373.1176	187.0625			P	1822.9480	911.9776	1805.9214	903.4644	18
5	432.1548	216.5810			460.1497	230.5785			S	1725.8952	863.4513	1708.8687	**854.9380**	17
6	560.2133	280.6103	543.1868	272.0970	588.2082	294.6078	571.1817	286.0945	Q	1638.8632	819.9352	1621.8367	811.4220	16
7	631.2504	316.1289	614.2239	307.6156	659.2454	330.1263	642.2188	321.6130	A	1510.8046	755.9060	1493.7781	747.3927	15
8	744.3345	372.6709	727.3080	364.1576	772.3294	386.6684	755.3029	378.1551	L	1439.7675	720.3874	1422.7410	711.8741	14
9	872.3931	436.7002	855.3665	428.1869	900.3880	450.6976	883.3615	442.1844	Q	1326.6834	663.8454	1309.6569	655.3321	13
10	985.4772	493.2422	968.4506	484.7289	1013.4721	507.2397	996.4455	498.7264	L	1198.6249	599.8161	**1181.5983**	591.3028	12
11	1056.5143	528.7608	1039.4877	520.2475	1084.5092	542.7582	1067.4826	534.2450	A	1085.5408	543.2740	1068.5143	534.7608	11
12	1113.5357	557.2715	1096.5092	548.7582	1141.5306	571.2690	1124.5041	562.7557	G	1014.5037	**507.7555**	997.4771	499.2422	10
13	1216.5449	608.7761	1199.5184	600.2628	**1244.5398**	622.7736	1227.5133	614.2603	C	957.4822	479.2448	940.4557	470.7315	9
14	1329.6290	665.3181	1312.6024	656.8049	**1357.6239**	679.3156	**1340.5973**	670.8023	L	**854.4730**	427.7402	837.4465	419.2269	8
15	1416.6610	708.8341	1399.6345	700.3209	**1444.6559**	722.8316	1427.6294	714.3183	S	**741.3890**	371.1981	724.3624	362.6849	7
16	1544.7196	772.8634	1527.6930	764.3502	**1572.7145**	786.8609	1555.6880	778.3476	Q	**654.3570**	327.6821	637.3304	319.1688	6
17	1657.8036	829.4055	1640.7771	820.8922	**1685.7986**	843.4029	1668.7720	834.8896	L	**526.2984**	263.6528			5
18	1794.8626	897.9349	1777.8360	889.4216	1822.8575	911.9324	1805.8309	903.4191	H	**413.2143**	207.1108			4
19	1881.8946	941.4509	1864.8680	932.9377	1909.8895	**955.4484**	1892.8630	946.9351	S	276.1554	138.5813			3
20	1938.9161	969.9617	1921.8895	961.4484	1966.9110	983.9591	1949.8844	975.4458	G	189.1234	95.0653			2
21									L	132.1019	66.5546			1

Error: try setting browser cache to automatic.

Error: try setting browser cache to automatic.

NCBI **BLAST** search of SSCPSQALQLAGCLSQLHSGL
(Parameters: blastp, nr protein database, expect=20000, no filter, PAM30)
Other BLAST web gateways

All matches to this query

Score	Mr(calc):	Delta	Sequence
43.0	2096.9983	0.0115	SSCPSQALQLAGCLSQLHSGL
33.1	2096.9983	0.0115	SSCPSQALQLAGCLSQLHSGL

Supplementary Information D - MS spectrum LLKCL-QLAGCL

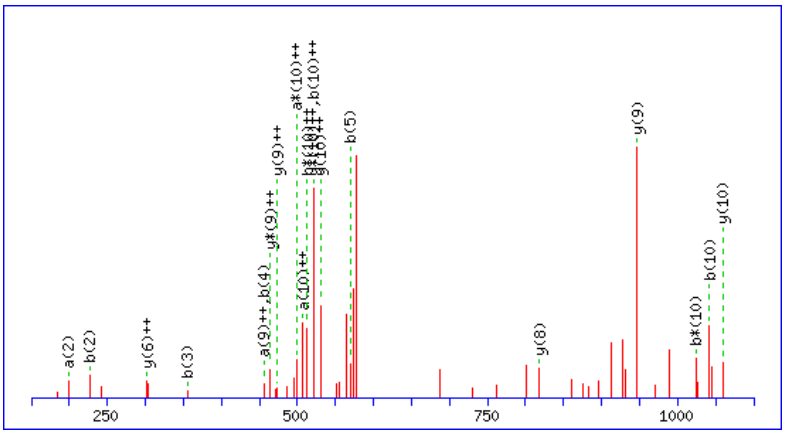
Mascot Search Results

Peptide View

MS/MS Fragmentation of **LLKCLQLAGCL**
Found in **P09919_ab_11_41**, a=LLKCL (15-19) cx b=QLAGCL (71-76) Granulocyte colony-stimulating factor OS=Homo sapiens GN=CSF3 PE=1 SV=1

Match to Query 162: 1171.648268 from(586.831410,2+) intensity(542379.3300)
Title: 516: Sum of 4 scans in range 2882 (rt=53.1124) to 2916 (rt=53.628) [E:\Data\RCPE1\RCPE1_II_20100917\Chymotr_08.RAW]
Data file E:\Data\RCPE1\RCPE1_II_20100917\Chymotr_08.RAW

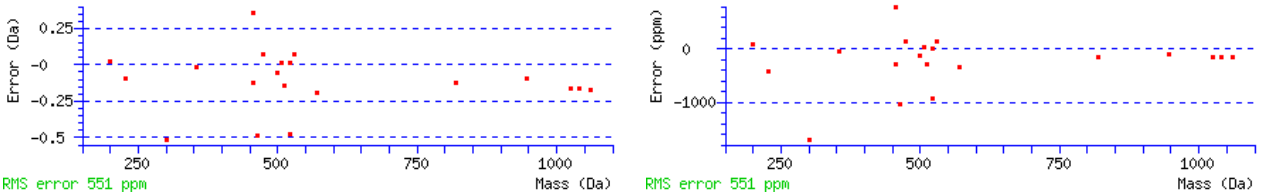
Click mouse within plot area to zoom in by factor of two about that point
Or, Plot from 150 to 1100 Da Full range



Monoisotopic mass of neutral peptide Mr(calc): 1171.6457

Variable modifications:
C4 : Disulfide bridge (C)
Ions Score: 20 **Expect:** 0.0011
Matches (Bold Red): 20/102 fragment ions using 42 most intense peaks

#	a	a^{++}	a*	a^{*++}	b	b^{++}	b*	b^{*++}	Seq.	y	y^{++}	y*	y^{*++}	#
1	86.0964	43.5519			114.0913	57.5493			L					11
2	**199.1805**	100.0939			**227.1754**	114.0913			L	**1059.5689**	**530.2881**	1042.5424	**521.7748**	10
3	327.2755	164.1414	310.2489	155.6281	**355.2704**	178.1388	338.2438	169.6255	K	**946.4849**	**473.7461**	929.4583	**465.2328**	9
4	428.2690	214.6381	411.2424	206.1249	**456.2639**	228.6356	439.2374	220.1223	C	**818.3899**	409.6986	801.3634	401.1853	8
5	541.3531	271.1802	524.3265	262.6669	**569.3480**	285.1776	552.3214	276.6644	L	717.3964	359.2018	700.3698	350.6886	7
6	669.4116	335.2095	652.3851	326.6962	697.4066	349.2069	680.3800	340.6936	Q	604.3123	**302.6598**	587.2858	294.1465	6
7	782.4957	391.7515	765.4692	383.2382	810.4906	405.7489	793.4641	397.2357	L	476.2537	238.6305			5
8	853.5328	427.2700	836.5063	418.7568	881.5277	441.2675	864.5012	432.7542	A	363.1697	182.0885			4
9	910.5543	**455.7808**	893.5277	447.2675	938.5492	469.7782	921.5226	461.2650	G	292.1326	146.5699			3
10	1013.5635	**507.2854**	996.5369	**498.7721**	**1041.5584**	**521.2828**	**1024.5318**	**512.7696**	C	235.1111	118.0592			2
11									L	132.1019	66.5546			1



NCBI **BLAST** search of LLKCLQLAGCL
(Parameters: blastp, nr protein database, expect=20000, no filter, PAM30)
Other BLAST web gateways

All matches to this query

Score	Mr(calc):	Delta	Sequence
19.6	1171.6457	0.0026	LLKCLQLAGCL
17.2	1171.6457	0.0026	LLKCLQLAGCL
14.1	1171.6457	0.0026	QLAGCLLLKCL
8.6	1171.6457	0.0026	QLAGCLLLKCL