Signal Peptide of Human Serum Albumin (residues 1-18) Forms Amyloid-like Aggregates

D. C. Thakur¹, Kumar Udit Saumya¹, Vipendra Kumar Singh¹, Rajanish Giri*¹

¹ Indian Institute of Technology Mandi, School of Biosciences and Bioengineering, VPO Kamand, Himachal Pradesh 175075, India

*Correspondence to: Dr. Rajanish Giri, School of Biosciences and Bioengineering, Indian Institute of Technology Mandi, Himachal Pradesh 175075, India.
Email: rajanishgiri@iitmandi.ac.in. Telephone number: 01905-267134, Fax number: 01905-267138

Abstract

Signal peptides (SPs) are essential tools in sorting or translocating the synthesized proteins. However, except few examples, their studies in isolation are quite limited. Here we asked a question of the aggregation propensity of signal peptide (residues 1-18) of Human Serum Albumin (HSA). Molecular dynamic simulations were used to acquire observations into the mechanism through which signal peptides of HSA self-aggregated in physiological conditions. Further, we confirmed our results by combining dye-based assays, atomic force microscopy, and transmission electron microscopy techniques. It was noted that the signal peptide of HSA (region 1-18) forms typical amyloid-like fibrils.

Keywords: Aggregation, Signal Peptides, Human Serum Albumin, Atomic Force Microscopy, High-Resolution Transmission Electron Microscopy.

Graphical
Introduction

Human serum albumin (HSA), the most abundant protein in the body, performs essential biological functions and binds to various ligands, including amino acids, fatty acids, ions, and hormones (PMID: 26055641; PMID: 30012492; PMID: 25674083). The half-life of HSA is approximately 20 days, and they are present in lymph, saliva, cerebrospinal, intestinal fluid, and plasma (Rabbani & Ahn, 2019; burek et al., 2011). Human serum albumin has worked as a molecular chaperone whose major function is restricting the aggregation of other proteins. They also protect enzymes from thermal degradation and aggregation (PMID: 20133158).

The signal peptides (SPs) are short peptides located at the N-terminal of proteins, involved in protein transport and release (PMID: 35118058). They are present in all eukaryotes and prokaryotes and have diverse applications in various industrial and scientific fields (PMID: 29958716). They are first identified by the signal recognition particle (SRP), a ribonucleoprotein complex, and subsequently enter between the twin transmembrane domain of Sec 61a (PMID: 9823886). In eukaryotes, the process starts when the SRP allows the N-terminal domain of a secretory protein. This episode permits co-translational shift of the protein into the lumen of endoplasmic reticulum (PMID: 11148214). Post transfer of the SPs into the lipid membrane, they are dissected by the membrane-bound SPs and subsequently break down by intermembrane proteases (PMID: 21810987).

The HSA signal peptide (residues 1-18) plays a crucial role in the folding and activity of recombinant proteins folded by 17 disulfide bonds (PMID: 29958716). Signal peptides are important molecular entities with a canonical function in targeted secretion translocation of proteins. They are also significantly involved in the production of recombinant proteins. Most of
the aggregation associated studies have been demonstrated the aggregation propensity of HSA protein in the presence of metal ions or alone (PMID: 30384272; PMID: 32023900). However, the misfolding and self-aggregating-like activity of signal peptides of HSA in isolation still pose unanswered questions (PMID: 29958716). Understanding the aggregation potential of signal peptides of HSA may provide valuable insights into the pathophysiology of related diseases. The formation of amyloid, a multi-step process that can lead to the production of highly toxic intermediates in biological systems, is associated with various diseases such as Alzheimer's, systemic amyloidoses, synucleinopathies, and type 2 diabetes mellitus (PMID: 30384272; PMID: 32023900; PMID: 19049981; PMID: 20079882; https://doi.org/10.1016/j.xcrp.2021.100599). Analyzing the aggregation propensity of the signal peptide of HSA can contribute to a better understanding of disease pathophysiology.

The aggregation of the signal peptide of HSA protein after its delivery to the intended destination raises questions about the consequences if the protein quality control system fails to degrade this cleaved-off signal peptide (PMID: 34050864). We aimed to investigate the fate of the signal peptide if it is not degraded by the protein homeostasis machinery or the protein quality control system and hypothesized that it may aggregate. Due to its solubility only in DMSO, studying the monomer was challenging. To test this hypothesis, we used in-silico tools, MD simulations, dye-based spectroscopy, and microscopy to predict and confirm the self-aggregation propensity of the HSA signal peptide (region 1-18) under physiological conditions.

2 Materials and Methods

2.1 Reagents
Signal peptide (residues 1-18) “MKWVTFISLLFLFSSAYS” of HSA protein was obtained from NCBI, chemically synthesized, and purchased from Genscript, USA, with 91.5% purity. TEM grid was acquired from TED PELLA INC., USA. Thioflavin T (ThT), Congo Red (CR), 8-Anilinonaphthalene-1-sulfonic acid (ANS), 1,1,3,3,3-hexafluoro-2-propanol (HFIP), were acquired from Sigma-Aldrich (St. Louis, MO, USA). Other chemicals used in this study were of the highest purity acquired from regional sources.

2.2 In silico prediction of fibril formation

The aggregation score (aggregation propensity region) of the signal peptide of HSA protein was predicted using different freely accessible data servers such as FoldAmyloid, Metamyb, TANGO, AGGRESCAN, and CamSol (Fernandez-Escamilla et al., 2004; Pawlicki et al., 2008; Conchillo-Solé et al., 2007; Sormanni et al., 2015; Garbuzynskiy et al., 2010). In short, parameters such as physiological setup (Temp-37°C and pH-7.4) were used to run the TANGO software. Whereas other databases, namely CamSol, Aggrescan Foldamyloid, and NHSA, were used in default settings. We used Peptide 2.0 (www.peptide2.com) to predict hydrophobicity, and the physiological parameters of the signal peptide were calculated with the help of the PROTPARAM database (www.expasy.org/protparam).

MD Simulation of HSA signal peptide (residues 1-18)

The simulations utilized the GROMACS software package (version 2022.3) along with the GROMOS96 54a7 force field (Schmid et al., 2019). The SPC (simple point charge) water model was used, and periodic boundary conditions (PBC) were implemented using a cubic simulation box. To ensure a minimum initial distance of 1 nm between proteins and the box boundaries, the system was set up accordingly. Nonbonded interactions were truncated at a 1.2 nm cutoff, and long-range electrostatics were computed using the Particle-mesh Ewald (PME) summation
method [DOI: 10.1021/acs.jpcb.5b04822]. The temperature was maintained at 300 K using a modified Berendsen thermostat (V-rescale), while the pressure was regulated at 1 bar with the Parrinello-Rahman barostat (Ng et al., 2021). Counterions were added to neutralize the system, and additional salt was included to achieve the desired concentration. Prior to the main simulations, energy minimization employing the steepest descent algorithm was conducted, followed by equilibration stages consisting of a 10 ns NVT simulation and a subsequent 10 ns NPT simulation. During equilibration, the protein backbone was restrained to maintain stability. These preliminary simulations were performed to generate the initial configurations necessary for the aggregation simulations. The starting conformation for the aggregation simulations was obtained from the 1 microsecond monomer simulations of HSA_SP (1-18). In aggregation simulation, six monomers were randomly placed in a periodic cubic box with an edge length of 10 nm, corresponding to a concentration of 0.15 M. The systems were solvated, and appropriate ions were added to neutralize the system and achieve the desired salt concentration. Before the production run, the aggregation simulations were further equilibrated using a 30 ns NVT simulation, followed by a 30 ns NPT simulation. The aggregation simulation was run for 300 ns. 

2.3 Thioflavin T assay

In vitro, aggregation was estimated using fluorescent dye-based assays such as ANS and ThT. A 200 μM ThT stock solution was prepared in distilled water, while a final concentration of 20 μM was used for a dye-based binding assay. Data was acquired using a black 96-well flat-bottomed plate through a microplate reader (TECAN Infinite M200 PRO). The excitation wavelength of 450 nm and emission spectra from 470 nm to 700 nm was used. Experiments were performed at least in triplicates.
2.4 ANS-dye-based fluorescence assay

ANS-dye-based binding analysis of signal peptide of HSA protein was conducted using a 20 μM concentration. A final concentration of 30 μM for both monomer and aggregated signal peptide of HSA protein was used in sodium phosphate buffer (20 mM; pH 7.4). Data were acquired in the TECAN Infinite M200 PRO microplate reader. The excitation wavelength of 380 nm and emission spectra ranging between 400 to 650 nm were used. The control used in the experiment was ANS dye only with buffer. Experiments were performed at least in triplicates.

2.5 Congo Red shift assay

Congo red (CR) dye is amyloid specific that binds with cross β-sheet rich amyloid-like aggregates and shows characteristic redshift in spectra. Mili-Q water was used to prepare a stock solution (0.1% w/v) of CR. Then CR solution with a concentration of 25 mM was added to the aggregated sample of the signal peptide of HSA protein (residues 1-18). Tecan Infinite M200 microplate reader was used to record the absorption spectra at 400-450 nm wavelength using a transparent 96-well plate. Experiments were performed at least in triplicates

2.6 High-Resolution Transmission Electron Microscopy (HRTEM).

For TEM analysis, 7 μL of the aggregate sample (diluted 10× in Milli-Q) was placed on 200 mesh grids of copper and incubated at RT for 60 s. The grid was rinsed softly, swirling in water droplets for at least 30 s three times afterward. An aqueous solution of ammonium molybdate (3% w/v) was used to stain the grids negatively. In the end, the grids were allowed to dry overnight, and then a snapshot was captured in HRTEM (FP 5022/22-Tecnai G2 20 S-TWIN, FEI) at 200kV accelerating voltage.

2.7 Atomic Force Microscopy (AFM).
The aggregate sample was first diluted 30 x times with Milli-Q and then drop casted using 7 μL on a newly dissected mica sheet. After 2 min incubation, the specimen was rinsed with deionized water and dried overnight in a desiccator. In the end, a snapshot was captured on AFM at the tapping mode (Dimension Icon from Bruker).

3. Results and Discussion

3.1 Prediction of aggregation propensity of the signal peptide (region 1-18) of HSA protein.

The aggregation tendency of every protein is modelled by evolution according to its actual presence in the organisms. HSA is an abundant protein in the human body, and it displays an excellent ligand-binding potential and is also a carrier of various exogenous and endogenous proteins. Protein aggregation is affected by several factors, such as the protein structure (primary, secondary, or tertiary), processing, and the environment of the protein in which protein exists. Aggregation or misfolding may reduce or entirely alter their biological characteristics. Therefore, protein aggregation remains a hurdle in the commercialization and development of biotechnology products. Various in-silico-based platforms that predict APR from the polypeptide sequence of signal peptides of HSA based on biochemical and physicochemical characteristics have been available. In the present study, we have used five web freely available high-precision server tools for APR prediction in the signal peptide of HSA (TANGO, AGGRESCAN, FoldAmyloid, CamSol, and AGGRESCAN). Our in silico server-based results revealed greater prediction scores for aggregation propensity of signal peptide region 1-18, especially the residue between the 2\textsuperscript{nd}-15\textsuperscript{th} (Figure 1 A-E). Human serum albumin involves transporting various exogenous and endogenous chemicals such as metals, drugs, amino acids, bile acids, toxic metabolites, hormones, and fatty acids (Taguchi et al., 2012; Otagiri & Chuang, 2009). Signal peptides are key tools that control the function and folding of recombinant proteins, which are
folded by disulfide bonds. Human serum albumin is a multifunctional protein with wide application in clinical settings.

The physicochemical properties of our signal peptide sequence were investigated utilizing Protparam and www.peptide2.com for the hydrophobicity index, respectively. We observed mostly hydrophobic amino acids, with an overall 11 of them allocated across the whole signal peptide of the HSA sequence (hydrophobicity index: 61.11%) (as shown in Figure 1).

**Figure 1:** Prediction of aggregation-rich regions in the signal peptide of human serum albumin.

Five web server tools predicted APRs of Human serum albumin’s signal peptide sequence (residues 1-18). (A) FoldAmyloid aggregation scan showing information on amyloidogenic region (B) MetAmyl server showing aggregation propensity as hotspot area profile value (C) TANGO server showing the percentage probability of the aggregation-prone region in the signal peptide of HSA. (D) Aggrescan servers showed the normalized hotspot area for APRs (E) CamSol solubility scan showed the solubility of the signal peptide of HSA.

### 3.2 Prediction of the aggregation of the signal peptide of HSA at the atomistic scale
Protein aggregation of signal peptides of HSA is a discussion of great concern to the researcher due to its involvement in protein homeostasis system and protein translocation. Numerous in-silico tools are available to predict aggregation propensity from analysis of amino-acid residues in a polypeptide chain. Similarly, investigation of the structural rearrangements during the aggregation and disaggregation process through numerous MD simulation techniques have been reported earlier. In this study, we therefore performed MD simulation of HSA signal peptide (region 1-18) to elucidate its aggregation behavior at the atomic level in presence of multiple monomeric subunits. GROMOS96 54a7 force field was applied to record the aggregation of signal peptides of HSA under physiological temperature. MD simulation was run for 300 ns, and a snapshot of signal peptides is illustrated in Figure 2. After the run of 50 ns, the signal peptides converted into dimers with following trimer, tetramer, pentamer, and at the last converted to hexamer (at 300 ns). So far, any study related to MD simulation that showed the propensity of self-aggregation behavior of signal peptides of HSA (region 1-18) under physiological conditions in not reported to date.
3.3 Binding of signal peptide sequence of HSA (residues 1-18) to amyloid-specific dyes and enhanced fluorescence.

Fluorescence is one of the most useful techniques for observing and structurally depicting amyloid fibrils or protein aggregation in pathogenic or clinical settings (Gasteiger et al., 2005; Hawe et al., 2008). The most common methods for quantifying and spotting amyloid-like fibrils depend on observing the fluorescence alterations in thioflavin T and 8-Anilinonaphthalene-1-sulfonic acid (ANS) dye and alteration in absorbance of the Congo red (CR), an azo dye. The ANS dye binds to the hydrophobic regions, enhancing fluorescence levels with a typical shift in emission spectra like a blue shift. First, we analyzed the fibril formation of the signal peptide of HSA protein by ANS-dye-based assay (as shown in Figure 2A). The reading was obtained at the
initial (0 h) and final (168 h) of the aggregate samples, demonstrating an increased fluorescence intensity of ANS approximately 1.5-fold. At 168 hrs, we observed a significant gain in fluorescence accompanied by a characteristic blue shift (549 nm to 516 nm).

In our study, we found that the signal peptide of the HSA protein shows a significant change in fluorescence signal when compared to monomer and ThT-only control. The aggregates of the signal peptide of HSA demonstrated a ~8-fold increment in ThT fluorescence intensity. It indicates the generation of amyloid-like fibrils in the aggregate-inducing environment under physiological settings, as shown in Figure (2B). We also noticed that CR binds to signal peptides of HSA proteins and demonstrates a significant shift in spectra from 488 to 491 nm (as shown in Figure 2C). The signal peptide of HSA (residues 1-18) was also positive for ANS assay, where we observed an increase in fluorescence spectra accompanied by characteristics of blue-shift in spectra from 541 nm - 498 nm.

![Figure 2](https://example.com/figure2.png)

**Figure 2:** Fluorescent dye binding assay for the analysis of signal peptides of HSA. (A) ANS fluorescence spectra of aggregated signal peptide sample, SP monomer, and ANS dye alone, respectively. The changes in spectra of the monomeric and aggregated samples are shown using an arrow and flecked line (B) ThT fluorescence spectra of the HSA (C) Congo Red signal peptide spectra of the aggregated signal peptide of HSA.
3.4 Signal peptides of HSA (residues 1-18) display fibril-like structure in morphology assessment of AFM and TEM.

Atomic force microscopy and High Resolution-TEM are versatile techniques for analyzing protein aggregation in vitro environments (Charnley et al., 2018; Sung et al., 2015). In our study, we analyzed and quantitatively computed the height of amyloid fibrils of our signal peptide samples (as shown in Figure 3A). The height of the fibrils of the signal peptide was approximately 20 nm [Figure 3A(III)]. Next, we confirmed the amyloid fibril formation of our aggregated sample by using transmission electron microscopy. A snapshot of the TEM image clearly showed a distinctive amyloid-like fibrillar structure with protruding short and long dense overlapped fibrils (as shown in Figure 3B).

**Figure 3:** Morphology analysis of signal peptide of human serum albumin amyloid fibrils. (A) AFM analysis represented the height (Y-axis) and width (X-axis) of aggregated signal peptide samples of HSA, where the mesh denoted the fibril used for measuring height. (B) HR-TEM snapshot image captured at 200 kV accelerating voltage.

4. Conclusions
From our initial prediction, we found the first 18 residues of HSA to have high aggregation propensity. We utilized the MD simulation to understand the self-aggregation nature of signal peptides of HSA (residues). Our MD simulation study clearly demonstrated the aggregation nature of signal peptides of HSA in the physiological environment. Further, our morphological analysis clearly verified the formation of typical amyloid-like fibrils. In the near future, it will help to understand the mechanisms of the entire HSA protein and its involvement in the transportation of various endogenous and exogenous substances. Our findings also indicate a possible role towards imbalance of HSA signal peptide in amyloidogenic diseases.

Credit authorship contribution statement

D.C. Thakur, Kumar Udit Saumya, Vipendra Kumar Singh: Data acquisition, analysis and interpretation of results, and manuscript writing. Rajanish Giri: Conceptualization, Project administration, Supervision, Validation, review & editing of the final manuscript.

Conflict of Interest

The authors declare no conflict of interest in the present work.

Acknowledgment

D.C. Thakur acknowledges the Indian Council of Medical Research Fellowship, New Delhi, India, and Kumar Udit Saumya acknowledges the Indian Institute of Technology Mandi for Ph.D. Fellowship. The authors thank the Indian Institute of Technology Mandi (Bio X, AMRC, and C4DFED Center) for the facilities. The authors highly acknowledge the National Supercomputing Mission (NSM) for providing computing resources of ‘PARAM Himalaya’ at IIT Mandi, which is implemented by C-DAC and supported by the Ministry of Electronics and
Information Technology (MeitY) and Department of Science and Technology (DST), Government of India.

References


