Supplement

Results

Modeling the protein structure of SP-H

An online BLAST search (http://blast.ncbi.nlm.nih.gov/) revealed that the only sequences with a high identity to SP-H in the UniProt (http://www.uniprot.org/) database are SP-H-homologs of other species, followed by a number very low score hits for short sections of putative regulatory proteins, mostly from different *Pseudomonas* species. An additional BLAST-search for similar sequences with already known 3D structure in the Protein Data Bank (http://www.pdb.org/) was also not successful. The identified hits had a too low sequence identity (< 20%) or coverage (only 20 of 94 residues). The sequence identity of SP-H to the already known SPs is also very low (ca. 10%). With that, no reliable templates were available for a classical homology modeling approach. Therefore, the sequence of SP-H was sent to the template-independent *ab initio* protein structure prediction server Robetta. Only this computationally expensive method was able to generate a 3D model which shows satisfying results in the model quality evaluation tools after minor optimizations. PROSA II produced a completely negative potential plot for all regions of the protein and the combined Z-score of the final model (-5.72) is comparable to the statistical average value for proteins of this length (-8.0). The stereochemical quality assessed by PROCHECK shows 94% of the amino acid dihedral angles in the most favored regions of the Ramachandran plot. These results indicate a good model quality and native-like fold of the final protein structure model.

To assess the protein model stability, a 20 ns molecular dynamics simulation was performed. As an indicator for model stability, the root mean square deviation (RMSD) plot of the protein backbone atoms was used. Already after 4 ns, it reaches a plateau where it remains stable until the end of the simulation (supplement Figure 2, black plot). No deterioration was observed in PROSA II or PROCHECK results after the MD simulation. This suggests that the obtained protein structure model for SP-H is stable enough for more sophisticated simulation studies.

As the result of the protein modeling, this is the first analysis that give an impression of the 3D structure of SP-H (Figure 1d). The most prominent structural features are a long and stable α-helix of the amino acids 7 to 31 and a β-sheet spanning the amino acids 55 to 73. Looking at the surface of the protein, no extensive hydrophobic regions for protein-lipid interactions are present. But there are single hydrophobic spots on the protein surface that are

formed only by a few amino acids (e.g. positions 28, 31, 34 or positions 88-91). However, no indications for transmembrane elements were found. The cysteine residues 45 and 56 could form a structure stabilizing intramolecular disulfide bridge. Since all three available cysteine residues are accessible on the protein surface, intermolecular disulfide bonds could be possible. Hence, an oligomerization of SP-H cannot be excluded based on the protein structure models.

Protein model with posttranslational modifications (PTMs)

Scanning the SP-H sequence for possible PTM sites with sequence based prediction tools gave the following results: NetNGlyc showed no potential *N*-glycosylation and NetAcet did not predict any acetylation. The NetOGlyc server suggests six threonines at the positions 55, 66, 69, 75, 76 and 93 to be modified with a GalNAc moiety, whereas the YinOYang prediction indicates a GlcNAc modification on Ser39, Thr76 and Ser78 with very high and on Ser82, 83 and 93 with low probability. NetPhos predicted seven phosphorylation sites, namely Ser32, Ser39, Thr55, Ser80, Ser82, Ser83 and Ser84, with the last four having a very high probability. Finally, the CSS-Palm server showed two of the three available cysteine residues (45 and 56) as potentially palmitoylated. In the case of multiple predictions for a position, the probability value was decisive. Furthermore, only surface accessible amino acids were modified. Finally, the protein modifications were manually added to the protein model according to Table 1.

The stability of the SP-H model with attached PTMs was examined by a 20ns MD simulation. The progress of the backbone atoms RMSD plot shows that the modified protein model is nearly as stable as the model without modifications (supplementary Figure 2, red plot). Until a simulation time of 11 ns, the RMSD values are almost the same. Thereafter, the plot for the modified model shows a higher fluctuation due to the influence of the large number of PTMs on the protein backbone. Nevertheless, no significant secondary structure changes or an unfolding of the protein can be observed. Analogous to the unmodified SP-H model, the modified protein structure can be considered as equilibrated and stable.

Where the modifications seem to have nearly no effect on the stability of the protein model in water, their influence on the protein properties becomes apparent. Depending on the modification type, polar (phosphorylation, glycosylation) or hydrophobic (palmitoylation) areas are formed on the protein surface, which could change the solubility of the protein in water and influence its interaction potential to a lipid environment significantly.

Materials and Methods

Tissues

The tissue samples were obtained from cadavers (5 male, 11 female, aged 33-76 years) donated to the Department of Anatomy and Cell Biology, Martin-Luther-University Halle-Wittenberg, Germany. The study was approved by the Institutional Review Board of the Martin Luther University Halle-Wittenberg in accordance with the Declaration of Helsinki. The used samples were dissected from the cadavers within a time-frame of 5-20 h postmortem. Previous to dissection, the history of each cadaver was studied. Samples that were affected by acute infections, tumors, recent trauma or surgical operations were not used in this study. Furthermore, all samples with a post-mortal interval longer than 20 hours were omitted. After dissection, half of the specimens were fixed in 4% paraformaldehyde for later paraffin embedding. The other half of the specimens were used for molecular-biological investigations and thus immediately frozen at -80°C. For the experimental part, we used lung, testis, heart, liver, kidney, parotid gland, umbilical cord, trophoblast, stomach, spleen and nasal mucosa samples.

Collection and analysis of the bronchoalveolar lavage (BAL)

BAL of the lung was isolated with 8 ml PBS for two times. BAL was collected and centrifuged at 1500 rpm for 5 min. The supernatants were frozen or subsequently analyzed by ELISA.

Collection and analysis of the sputum from cystic fibrosis (CF) patient

The sputum was transferred directly into the tube, the specimen were immediately frozen on dry ice. All the samples were then transported on dry ice from the clinic to the lab within 2–6 h and then immediately stored at −80°C until analyzed. The sputum was obtained from CF patients with chronic pulmonary colonization by P. aeruginosa and other bacteria. The samples were collected and centrifuged at 1500 rpm for 5 min. The supernatants were frozen or subsequently analyzed by ELISA. The CF patients were grouped according to their lung function (FEV_1 in % age predicted). Patients had a FEV_1 value lower than 60% ($FEV_1 < 60\%$), patients higher than 60% ($FEV_1 > 60\%$).

Cloning and protein expression of SFTA3 using E. coli

For protein expression in *Escherichia coli*, the coding region without signal sequence was cloned into the pBAD vector containing a 6xHis tag using the pBad Topo TA expression kit. Instructions were provided by the pBAD TOPO TA expression kit (Invitrogen Life Technologies, Carlsbad, CA). For the amplification of SP-H we used the following primers: sense 5'-ATG AGA GCC GGG TTT TCT GAC-3' and antisense 5'-TGC AGT ATG AAT AAT TAA CAT C-3'. The PCR products were ligated into the TOPO/TA vector and transformed into TOP100 E. coli cells (Invitrogen, Carlsbad, CA). Positive colonies containing the inserted gene in the proper orientation were identified using PCR. For recombinant expression of SP-H, a system carrying an inducible T7 promoter and an N-terminal 6xHis tag was used. Overnight cultures of LB broth (5 ml) containing 0.02 mg/ml ampicillin (amp+) were used to inoculate the 50 ml LB/amp+ cultures. The cultures were induced with L-Arabinose in midlog phase (OD_{600} 0,8 – 1) and incubated at 37°C for 4 h. The cells were harvested by centrifugation and the supernatant was discarded. The cell pellet was resuspended in 20 mM Tris, 500 mM NaCl, and 5 mM imidazole with a pH of 7.9 (1/20 of the initial culture volume). The resuspended cells were placed on ice and lysed by ultrasonication. The soluble protein of the cell lysate was isolated by centrifugation at 13.000 rpm for 20 min at 4°C. The supernatant was use directly for Antibody testing.

RNA Preparation and cDNA Synthesis

For conventional reverse transcriptase polymerase chain reaction (RT-PCR), tissue biopsies were crushed in an agate mortar under liquid nitrogen and homogenized (Polytron, Norcross, GA). Total RNA was extracted from the tissue biopsies by RNeasy Mini Kit; Qiagen, Hilden, Germany.

Total RNA was extracted from cultured A549 cells (PeqGold reagent; PeQLab, Erlangen, Germany). Crude RNA was purified with isopropanol and repeated ethanol precipitation, and contaminated DNA was destroyed by digestion with RNase-free DNase I (30 minutes 37°C; Boehringer, Mannheim, Germany). The DNase was heat-inactivated for 10 min at 65°C.

Reverse transcription of all RNA samples to first-strand cDNA (RevertAid H Minus M-MuL V Reverse Transcriptase Kit; Fermentas, St. Leon-Rot, Germany) was performed according to the manufacturer's protocol. Two micrograms total RNA and 10 pmol Oligo $(dT)_{18}$ primer (Fermentas) were used for each reaction. The ubiquitously expressed β-actin and HPRT which proved amplifiable in each case with the specific primer pair, served as the internal control for the integrity of the translated cDNA.

Polymerase Chain Reaction (PCR)

For conventional PCR, we used conditions as previously described with the following primers: SP-H 1 sense 5'- CAC CAT GAG AGC CGG GT -3', antisense 5'-TCA TGC AGT ATG AAT AAT TAA CAT CTT-3' (ca. 275 bp) and SP-H 3 sense 5'- GAG AGG GAA AAG CGG ATA CC -3', antisense 5'- CGG ATG AAC TCC TGC TTT GT -3' (ca. 112 bp) [9]. For verification and comparison, bacterial plasmids carrying the genes for the investigated protein were used as a reference (German Resource Centre for Genome Research GmbH; SP-H: IRAKp961P2288Q). PCR products were also confirmed by BigDye sequencing (Applied Biosystems, Foster City, CA). To estimate the amount of amplified PCR product, we performed a ß-actin PCR with specific primers (sense 5'- CAA GAG ATG GCC ACG GCT GCT-3', antisense 5'- TCC TTC TGC ATC CTG TCG GCA-3', 275 bp) for each investigated tissue. PCR products were also confirmed by BigDye sequencing (Applied Biosystems, Foster City, CA).

Quantitative Real-Time RT-PCR

Samples were analyzed by CFX96 Touch Real-Time PCR Detection System (Bio-Rad Laboratories). Real-time RT-PCR was performed with SP-H primers (see above) to allow calculation of the relative abundance of transcripts. The PCR reaction contained 10 µL SsoFast EvaGreen Supermix (Bio-Rad Laboratories, München, Germany), 0.4 µL 10 pmol of each primer mix, and 1 µL of each cDNA in a final volume of 20 µL. In each plate qPCR was performed with a cycle of 2 min 98°C, 50 cycles at 5 s 95°C, 10 s 60°C, followed by 5 s 65°C and 5 s 95°C, to confirm amplification of specific transcripts. A standard curve was generated by fourfold serial dilutions of cDNA from non stimulated cells. To standardize mRNA concentration, the transcript levels of the housekeeping gene small ribosomal subunit

(hHPRT) were determined in parallel for each sample, and relative transcript levels were corrected by normalization based on the hHPRT transcript levels.
The primers for hHPRT were as follows: forward 5′- TGA CAC TGG CAA AAC AAT GCA -3′ and reverse 5′- GGT CCT TTT CAC CAG CAA GCT -3′. All real-time RT-PCR's were performed in triplicate, and the changes in gene expression were calculated by the $^{\Delta\Delta}C_t$ method.

Western blot analysis

For Western blots, lung tissue (standardized ratio: 100mg wet weight/400µm buffer containing 1% SDS and 4% 2-mercaptoethanol) was extracted as previously described [10]. The protein was measured with a protein assay based on the Bradford dye-binding procedure (BioRad, Hercules, CA). The total protein (30µg) was then analyzed by Western blot. Proteins were resolved by reducing 15% SDS-polyacrylamide gel electrophoresis, electrophoretically transferred at room temperature for 1 h at 0.8 mA/cm^2 onto 0.1 µm pore size nitrocellulose membranes and fixed with 0.2% glutaraldehyde in phosphate-buffered saline for 30min. Bands were detected with primary antibody to SP-H (1:250) and secondary antibody (anti-rabbit IgG, respectively, conjugated to horseradish peroxidase, 1:5.000) using chemiluminescence (ECL- Plus; Amersham-Pharmacia, Uppsala, Sweden). Human lung was used as the control. The molecular weights of the detected protein bands were estimated using standard proteins (Prestained Protein Ladder, Fermentas, St. Leon-Rot, Germany) ranging from 10 to 170 kDa.

Immunohistochemistry

For immunohistochemistry, tissue specimens from healthy tissues of cadavers were embedded in paraffin, sectioned (6µm) and dewaxed. Immunohistochemical staining was performed with the polyclonal antibody against SP-H. Antigen retrieval was performed by microwave pretreatment for 10 min and non-specific binding was inhibited by incubation with porcine normal serum (Dako) 1:5 in Tris-buffered saline (TBS). Each primary antibody (1:50 - 1:100) was applied overnight at room temperature. The secondary antibodies (1:300) were incubated at room temperature for at least 4 h. Visualization was achieved with aminoethylcarbazole (AEC) for at least 5 min. Red stained areas within the tissues indicate a positive antibody reaction. After counterstaining with hemalum, the sections were mounted in

Aquatex (Boehringer, Mannheim, Germany). Two negative control sections were used in each case: one was incubated with the secondary antibody only, and the other one with the primary antibody only. The slides were examined with a Keyence Biorevo BZ9000 microscope.

A549 cells cultured on glass coverslips were washed three times in TBS (0.14 M NaCl in 20 mM TRIS/NaCl buffer, ph 7.4) and fixed in −20°C cold methanol for 5 min. Following washing three times in TBS, cells were incubated with 5 mg/ml hyaluronidase for 30 min at 37°C in a humid chamber. Afterwards, cells were incubated over night with the polyclonal antibody against SP-H at 4°C. For preparation of a negative control, the primary antibody was replaced by nonimmune serum. As secondary antibodies, respective fluorescein-isothiocyanate (FITC)-conjugated antibodies (Alexa 488, green) diluted 1:200 with TBS were used. Cells were embedded with DAPI-Glycerol (PBS-Glycerol 1:1, by adding 10 µl of 2mg/ml DAPI stock solution) on glass slides. The slides were examined with a Keyence Biorevo BZ9000 microscope.

Immunoelectron Transmisson Electron Microscopy

Bronchial tissue was fixed with 4% PFA. After several washes with PBS, the tissue was treated with the primary antibody 1:50-100. Incubation was done overnight at 4° C. Then, tissue was washed repeatedly with PBS, followed by incubation with gold-labeled antibodies for 2 hours. The samples were washed with PBS followed by incubation with 2.5% glutaraldehyde. The samples were incubated with 0.5% osmium tetroxide and after washout with PBS. Then they were treated with silver enhancement. The tissue was washed 2 times with distilled water and mixed with 4% low-melt agarose until the agarose solidified. The agarose block containing the homogenously distributed bacteria was stored for 2 days in 70% ethanol.

After dehydration in graded concentration of ethanol, the tissue block was incubated twice 1:1 in a 100% ethanol: acetone mixture and once in 100% acetone. Finally the lung tissue block was infiltrated with increasing concentrations of Epon with the following acetone-Epon-mixtures: 1:3 Epon, 2:3 acetone and 2:3 Epon, 1:3 acetone and Epon 100% without acetone. The embedded block was polymerized at 60°C for 24 h

and 90°C for 48 h. All TEM investigations were performed with a Zeiss TEM 902 ESI (Carl Zeiss, Jena, Germany). The black dots indicate positive antibody reactivity.

Process for cutting ultrathin sections

Semithin sections of embedded tissue were examined. The 50 nm sections were trimmed with a diamond knife. The sections were taken up from water with 200-mesh copper, dried and counterstained with aqueous uranyl-acetate and lead citrate to increase the contrast for transmission electron microscopy. The grids were examined with a transmission electron microscope (Zeiss 900).

Cell Culture

Human Lung Epithelial Carcinoma cells A549 were cultured in Dulbecco's modified Eagle's medium (DMEM/HAMs F12 1:1; PAA Laboratories GmbH, Pasching, Austria) containing 10% fetal calf serum (FCS; Biochrom AG, Berlin, Germany) in a humidified 5% CO_2 incubator at 37°C. For stimulation experiments, the A549 cells (1×10^6cells) were grown at confluence and before treatment, the cells were washed in phosphate-buffered saline (PBS) and incubated in serum-free medium overnight. They were then treated either with interleukin (IL)-1β (50 ng/mL; ImmunoTools, Friesoythe, Germany), or (IL)-23 (50 ng/mL; ImmunoTools) or both together and lipopolysaccharide (LPS, 20 ng/mL; Nürnberg, Germany) in medium containing serum for 24 hours. All experimental procedures were performed under normoxic conditions [31]. On completion of each experiment cell were incubated in PeQGold andstored and -80°C until they were processed for RNA extraction.

Enzyme-Linked Immunosorbent Assay (ELISA)

The ELISA analysis was performed using kits and the regarding protocols from USCN Life Science Inc. Wuhan. By comparing with the standard series and the determined values for antigen concentration (protein concentration), each sample was calculated in ng/mg.

Protein structure model creation

To obtain the 3D model of the protein structure, different approaches were used. Initially, the implementation of homology modeling in YASARA [32, 33] was performed. Furthermore,

the protein sequence was sent to the online servers I-TASSER [34] and LOOPP (http://cbsuapps.tc.cornell.edu/loopp.aspx), which use the so called threading-approach. Additionally, the sequence was also submitted to the online *ab initio* folding server Robetta [35]. Only this computationally very expensive method led to the final protein structure model for SP-H, with which energy minimizations and MD refinements [32] in YASARA were performed to further improve the intramolecular interactions and stereochemistry. The stereochemical quality was evaluated by PROCHECK [36] and PROSA II [37] was used to assess the quality of the entire protein fold or identify possibly misfolded regions of the protein model. PROSA II contains knowledge based mean fields derived from statistical analysis of well resolved protein X-ray structures. Both validation programs can give clear hints if the structure model resembles a native-like fold. The final SP-H model was accepted by and deposited at the Protein Model DataBase PMDB (http://mi.caspur.it/PMDB/) and received the PMDB id PM0079092 for free download.

To check the stability of the final model, a 20 ns MD simulation was performed in YASARA. The MD was done in a water box with a physiological NaCl concentration of 0.9% and the YASARA2 force field [32].

Prediction of protein modifications

Statistical prediction tools were used to analyze SP-H *in silico* for posttranslational modifications based on the amino acid sequence. Therefore, different programs were used which are all collected on the ExPASy bioinformatics resource portal (http://www.expasy.org/). The protein sequence was scanned for acetylation, *N*-glycosylation, *O*-glycosylation with N-Acetylglucosamine (GlcNAc) or N-Acetylgalactosamine (GalNAc) and phosphorylation with NetAcet, NetNGlyc, NetOGlyc, YinOYang and NetPhos, respectively. Furthermore, the possibility of palmitoyl chains bound to the three available cysteine residues of SP-H was checked by CSS-Palm. Predicted modifications were added manually to the protein structure model, followed by an energy minimization in YASARA. The resulting modified SP-H model was accepted by and deposited at the Protein Model DataBase PMDB and received the PMDB id PM0079093 for free download. The stability of the modified 3D model was checked by a 20 ns MD simulation in YASARA similar to the calculation for the unmodified structure model (water box, 0.9% NaCl, YASARA2 force field).

Molecular dynamics simulations

To study possible interactions between protein model and lipid environment, the simulation system should be as close as possible to the native state. For the MD simulations performed in this work, the major component of the lung surfactant [6] was chosen as lipid part, which is dipalmitoylphosphatidylcholine (DPPC). To meet the present picture of the lung surfactant lipid system [38], the DPPC molecules were arranged as a monolayer patch with the polar head groups facing a liquid phase and the alkyl chains facing the air.

The protein-lipid simulations were performed with the GROMACS package version 4.5.4 [39]. The united-atom G53a6 force field [40] was modified after Kukol [41] to produce reasonable data for a DPPC-lipid system. To assure the proper simulation of the modified protein model, the force field was extended by building blocks for phosphorylated serine and threonine, palmitoylated cysteine and serine or threonine residues which are *O*-glycosylated with GlcNAc or GalNAc. Parameters for all these groups were taken from building blocks of the original G53a6 force field and in the case of the phosphorylated amino acids from the G43a1p force field (GROMACS user contribution of G. R. Smith, available from http://www.gromacs.org). The CELLmicrocosmos MembraneEditor 2.2 [42] was used to build the fundamental simulation system layout. Each of the two distinct DPPC monolayers consists of 128 lipid molecules. On the polar head group side, the monolayers are separated by a water phase. On the side of the lipid alkyl chains, they are divided by a vacuum phase due to the periodic boundary conditions applied in all three dimensions. For the simulation start, one copy of the protein model was placed inside the water phase between the two lipid layers and was neutralized with Na^+ or Cl^- ions, if necessary. The starting orientation of the protein model was altered slightly for different simulations. The total size of the resulting simulation systems was approximately 60.000 atoms.

After an initial equilibration period (500 ps), the simulations were carried out for 50 ns with the Nosé-Hoover thermostat [43, 44] at 323 K and the Parrinello-Rahman barostat [45] with semi-isotropic coupling and a reference pressure of 1 bar. The compressibility of the system was set to 0 in z direction to conserve the simulation setup and stabilize the system. The LINCS constraint algorithm [46] was used to fix the bond stretching of all hydrogen involving atom bonds, allowing a simulation time step of 2 fs. Electrostatic interactions were calculated with the Particle Mesh Ewald (PME) algorithm [47] as implemented in GROMACS with the cutoff at 1.2 nm and the van der Waals potential switched off between 1.2 and 1.3 nm. The neighbor list was updated every 10 steps and no dispersion correction

was applied. The analysis of the simulations was done with the tools integrated in the GROMACS package. Visualization of the structures and trajectories was done with VMD [48] and YASARA.

Statistical analysis

Differences were evaluated for significance by the Student's two-tailed t test for parametric data. * $p \leq 0{,}05$, ** $p \leq 0{,}01$, *** $p \leq 0{,}001$. Data are given as mean values ± s.e.m.

Reference

32. Smith BT. Cell line A549: a model system for the study of alveolar type II cell function. *Am Rev Respir Dis* 1977: 115(2): 285-293.

33. Krieger E, Joo K, Lee J, Raman S, Thompson J, Tyka M, Baker D, Karplus K. Improving physical realism, stereochemistry, and side-chain accuracy in homology modeling: Four approaches that performed well in CASP8. *Proteins* 2009: 77 Suppl 9: 114-122.

34. Krieger E, Darden T, Nabuurs SB, Finkelstein A, Vriend G. Making optimal use of empirical energy functions: force-field parameterization in crystal space. *Proteins* 2004: 57(4): 678-683.

35. Roy A, Kucukural A, Zhang Y. I-TASSER: a unified platform for automated protein structure and function prediction. *Nat Protoc* 2010: 5(4): 725-738.

36. Kim DE, Chivian D, Baker D. Protein structure prediction and analysis using the Robetta server. *Nucleic Acids Res* 2004: 32(Web Server issue): W526-531.

37. Laskowski RA, MacArthur DS, Moss DS, Thornton JM. PROCHECK: a program to check the stereochemical quality of protein structures. . *J Appl Cryst* 1993: 26: 283–291.

38. Sippl MJ. Recognition of errors in three-dimensional structures of proteins. *Proteins* 1993: 17(4): 355-362.

39. Alonso C, Alig T, Yoon J, Bringezu F, Warriner H, Zasadzinski JA. More than a monolayer: relating lung surfactant structure and mechanics to composition. *Biophys J* 2004: 87(6): 4188-4202.

40. Hess B, Kutzner C, van der Spoel D, Lindahl E. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J Chem Theory Comput* 2008: 4: 435–447.

41. Oostenbrink C, Villa A, Mark AE, van Gunsteren WF. A biomolecular force field based on the free enthalpy of hydration and solvation: the GROMOS force-field parameter sets 53A5 and 53A6. *J Comput Chem* 2004: 25(13): 1656-1676.

42. Kukol A. Lipid Models for United-Atom Molecular Dynamics Simulations of Proteins. *J Chem Theory Comput* 2009: 5: 615–626.

43. Sommer B, Dingersen T, Gamroth C, Schneider SE, Rubert S, Kruger J, Dietz KJ. CELLmicrocosmos 2.2 MembraneEditor: A Modular Interactive Shape-Based Software Approach To Solve Heterogeneous Membrane Packing Problems. *J Chem Inf Model* 2011.

44. Hoover WG. Canonical dynamics: Equilibrium phase-space distributions. *Phys Rev A* 1985: 31(3): 1695-1697.

45. Nose´ S. A molecular dynamics method for simulations in the canonical ensemble. *Mol Phys* 1984: 52: 255–268.

46. Parrinello M, Rahman A. Polymorphic transitions in single crystals: A new molecular dynamics method. . *J Appl Phys* 1981: 52: 7182–7190.

47. Hess B, Bekker H, Berendsen HJC, Fraaije JGEM. LINCS: A linear constraint solver for molecular simulations. *J Comput Chem* 1997: 18: 1463–1472.

48. Essmann U, Perera L, Berkowitz ML, Darden T, Lee H, Pedersen LG. A smooth particle mesh Ewald method. *Journal of Chemical Physics* 1995: 103, 8577–8593.

49. Humphrey W, Dalke A, Schulten K. VMD: visual molecular dynamics. *J Mol Graph* 1996: 14(1): 33-38, 27-38.

Supplement figure legends

Supplement figure 1:

Western blot analysis of recombinant expressed SP-H protein using *E. coli* BL21. rb a SP-H) anti-SP-H-antibody revealing distinct proteins at expected molecular weights at 18 kDa and 43 kDa. **maV5)** anti-V5-antibody revealing bands at the expected molecular weight of 18 kDa and 43 kDa.

Supplement figure 2:

Root-mean-square deviation (RMSD) plot for the SP-H backbone atoms. The RMSD was calculated over the whole simulation process to check the stability of the unmodified (black) and posttranslationally modified (red) protein structure model. Rare and small plotfluctuations indicate a stable model.