Access to Higher Order Structure Information of Microcrystalline Biopharmaceuticals by Solid-State NMR spectroscopy in its formulated state

Soumya Ranjan Pujahari, Ashutosh Kumar*

Department of Biosciences and Bioengineering, IIT Bombay, 400076

Introduction

Low-volume, high-dose delivery of biotherapeutics is an emerging approach adopted by biopharmaceutical industries [1]. The monoclonal antibody is a significant class of biotherapeutics for treating many life-threatening diseases such as cancer, neurodegenerative diseases, diabetes, arthritis, etc., requiring high doses (>100 mg/mL) for the targeted delivery and optimum efficacy. Such highly concentrated drugs are formulated so that the lowest possible volume of the drug can be administered [2]. However, high doses of biotherapeutics are more prone to physical and chemical instability such as aggregation [3], self-association [4,5], liquid-liquid phase separation [6–8], oxidation, and fragmentation [9,10]of the sample. Protein crystallization is the best possible approach to overcome such problems, which is less viscous than the solution formulation [11–13] of similar concentration [14], less prone to degradation, and enhanced delivery pharmacokinetics over the solution formulation. Crystallization is also helpful in isolating and purifying proteins, enhancing biopharmaceutical manufacturability[15,16]. The use of drugs in microcrystalline suspension form enhances their bioavailability and allows for sustained release compared to other dosage forms [17–19].Insulin is currently the only biological product marketed as a micro-crystalline (NPH formulation) and biphasic suspension formulation (mixture of suspension and soluble formulation). However, the crystallization of full-length monoclonal Antibodies appropriate for batch-scale manufacture and suitability for human administration has recently been a hot topic of research [20,21]. The HOS of insulin has been studied both in solution [22] and crystalline[23,24] form and it was concluded that insulin exists in a hexameric form stabilized by two zinc (zn^{2+}) ions and a phenolic compound (i.e., m-cresol) [23] present in the microcrystals. Upon injection at the subcutaneous site, while the solution insulin quickly goes into the bloodstream, the microcrystals of the suspension formulation form a concentrated heap at the site of injection [25,26]. The microcrystalline heap, upon dissolution, releases soluble insulin as hexamers, which in turn dissociates into dimers and monomers, each with a different dissociation constant, facilitating controlled release [27]. Hence, the HOS of Insulin, crystal morphology, crystal packing, crystallinity, and composition of the crystal are essential factors that can potentially impact the action profile of the suspension formulation [27]. As the formulation of the drug product was carried out in its microcrystalline state, investigation of the HOS of the protein (drug) in the native state (crystalline form) is of utmost importance.

Protein higher order structure (HOS) analysis is critical to understanding the molecular structure and defining a biologics critical quality attributes (CQAs). According to the International Conference on Harmonisation(ICH) Q6B guidelines, it is also essential for the Test Procedures and Acceptance Criteria for Biological Products [28]. Alteration in the higher-order structure of the protein can impact the stability, quality, safety, and efficacy of biopharmaceutical products, leading to an increased potential for immunogenicity and loss of

biological function. Therefore, appropriate analytical tools are required to measure and monitor secondary, tertiary, and quaternary protein structure during early and late phase characterization and as part of comparability studies in the native formulated state. Several biophysical techniques are used regularly for biopharmaceutical characterization. For example, differential scanning calorimetry (DSC) [29] is used to monitor the phase transition of biomolecules, and Fourier transform infrared spectroscopy (FT-IR) [30] and Circular Dichroism (CD) [31] are used to decipher the secondary structure of proteins. Whereas H/D exchange mass spectrometry (HDX-MS) [32] is helpful for peptide/protein fragment mass determination, X-ray diffraction [29,33], liquid and solid state Nuclear Magnetic Resonance (NMR) [34,35], and cryo-electron microscope (cryo-EM), [36] can elucidate three-dimensional global and local structure determination as well as drug-receptor interactions. Barring liquid state NMR, biophysical techniques discussed above involve sample manipulation and may miss out on the native structure in its formulated state. NMR has several advantages over the others as it offers high structural resolution without any sample manipulation [37–39] and therefore is highly suited for characterizing biopharmaceuticals.

Solid-state Nuclear Magnetic Resonance (ssNMR) spectroscopy plays an essential role in the field of small molecule pharmaceutical analysis, as nearly 80% of formulated drug products are available in different solid forms. ssNMR provides valuable atomic-level information throughout the drug development process. However, the commonly used ^{13}C detection methods for studying pharmaceutical products face challenges due to the low natural abundance of this isotope, resulting in less sensitive experiments and longer acquisition times. Lately, with the development of ultra-fast magic-angle spinning (MAS) NMR probes, it is now possible to incorporate ^{1}H detection schemes in solid-state NMR [40]. This significantly benefits from its rich abundance and high gyromagnetic ratio to get the sensitivity enhancement necessary for natural abundance samples. These advances allowed the extension of the ssNMR application to suspension formulations of therapeutic proteins [41]. In pharmaceutical characterization, many established 2D methods employing proton detection primarily focus on ^{1}H-^{1}H, ^{15}N-^{1}H, and ^{13}C-^{1}H correlation spectra, serving as distinctive patterns for identifying and analysing pharmaceutical substances [41,42]. It is worth mentioning that Proton-detected solid-state NMR (ssNMR) is also used for determining the structure of amyloid fibrils however the samples tend to have broadened line widths due to their inherent inhomogeneity [43]. Nevertheless, there are notable studies in this area; Daskalov et al. [44] demonstrated the high-resolution structure of HELLF fibrils, while Bahri et al. [45] revealed both close-contact and long-distance interactions within Ab (1–42) fibrils exclusively through proton detection. Shi et al. [46] investigated the helical structure of the cytoskeletal bactofilin BacA where proton detection was used. However, it should be noted that these samples were isotopically labelled for resonance assignment and ^{1}H-^{1}H distances were used for structure calculations.

The side chains of amino acids play a crucial role in maintaining the three-dimensional structure of proteins and the arrangement of subunits in protein complexes. Since the folding and shape of the protein directly impact the functionality of a protein drug, side chains can provide valuable insights into the structural integrity. The latest guidelines for developing biosimilars strongly recommend reporting on the folding and aggregation properties, as well as characterizing higher-order structures (HOS). Methyl groups, in particular, have been

extensively used in biomolecular NMR as highly sensitive site-specific probes for structure [47–50], dynamics [51,52], and interaction studies [53–55]. Due to their well-dispersed nature in the protein's 3D structure, methyl chemical shifts can provide valuable information on subtle alterations inside the hydrophobic pocket of the protein that can alter the native folding as well as potential changes in the oligomerization state of a protein. Alternatively, these can also be explored for identifying binding epitopes in the protein receptor interactions [56].

Objectives

a) In this study, we have demonstrated an NMR-based approach that can be used for comparative assessment of higher order structure for biphasic microcrystalline suspension biotherapeutics in its formulated state. Biphasic microcrystalline formulations are a combination of regular formulations (with soluble components) and microcrystalline proteins (insoluble components). Due to the colloidal biphasic nature, its characterization by established liquid-state NMR methods is not feasible, so we have used a combinatorial approach of liquid and solid-state NMR to investigate the biphasic suspension formulation. In our study, we separated the biphasic Insulin drug into two fractions (solid and soluble) by centrifugation and carried out 1D ^{1}H liquid NMR and 1D ^{13}C CP-MAS NMR experiments for the soluble and solid fraction, respectively. For comparative assessment, the spectral outcomes from these experiments were further assessed using chemometric tools - Principal Component Analysis (PCA) and Mahalanobis distance (D_M). In this study, we have used four batches each of marketed biphasic microcrystalline Insulin suspension (Insulin 70/30 and Insulin 50/50) available in the Indian market branded as Huminsulin® and Insugen®. Eventually, we demonstrated that a combination of simple NMR methods and chemometric assessment could provide a high-resolution assessment of biphasic pharmaceutical suspensions that are not easily amenable to structural characterization in their native formulation states.

b) We also attempted to compare the spectral data of the ^{1}H-detected ^{13}C-^{1}H HETCOR experiment; however, we encountered limitations in achieving sufficient resolution. To solve this problem, we introduced a novel approach for solid-state NMR analysis of pharmaceutical suspensions at fast magic-angle spinning (MAS), specifically targeting the methyl and aliphatic side-chain characteristics. The proposed method incorporates a unique band selective excitation scheme and combines proton detection with non-uniform sampling (NUS) to reduce the time required for spectrum acquisition. The reconstruction of NUS data using a compressed sensing algorithm further enhances resolution, allowing for site-specific characterization of the higher-order structure of insulin in its microcrystalline suspensions. In our study, we also utilized a similar machine-learning platform called SHIFTX2, which combines ensemble machine-learning methods with sequence alignment-based approaches to calculate protein chemical shifts for both backbone and side chain atoms [57]. The predicted chemical shifts were instrumental in the assignment of a few key resonances in two-dimensional ^{13}C-^{1}H HETCOR spectra. Insulin, in its drug-formulated state, exists in a thermodynamically stable hexameric state, exhibiting conformational dynamics between R6, T6, and T3R3 states. With substantially

enhanced resolution and SHIFTX2 predicted assignments, we were able to deduce the diverse range of conformational populations and elucidate the potential mechanism of interconversion between these conformational states.

Material and Methods

Materials

All Insulin microcrystalline biphasic drugs were purchased from the market. 10 mL multiple-dose biphasic vial of each of Huminsulin® 70/30 (100IU/ml) and Huminsulin® 50/50 (40IU/ml) (* 100IU is not available), Insugen® 70/30 (100IU/ml) and Insugen® 50/50(100IU/ml) was used for the study (100 IU ~ 3.5 mg equivalent of Insulin). Biphasic Insulin drugs are mixture of regular insulin (water soluble) and microcrystalline (Neutral Protamine Hagedorn) formulation in different ratios (In 70/30, 70% microcrystalline and 30% soluble Insulin; similar in 50/50 both are in equal ratios). 10 mL multiple-dose clear formulation of each Huminsulin® R (100IU/ml) and Insugen® R (100IU/ml) and microcrystalline (NPH formulation) of Huminsulin® N was purchased over the counter [All Huminsulin® products are from Eli Lilly (Indian marketed) and Insugen® products are from Biocon Biologics Limited, Bengaluru, India]. Huminsulin® and Insugen® are independently approved generic drugs with similar Insulin formulations with minuscule differences in the excipients. Pierce™ BCA kit was purchased from Thermo Fischer Scientific, India. Uniformly ^{13}C, ^{15}N labeled microcrystalline SH3 domain protein was purchased from Cambridge isotope Laboratories, Inc. Humulin® 70/30 (mixture of regular human insulin, Humulin® R and the intermediate action of Humulin® N isophane suspension) and Humulin® N (isophane insulin human suspension) by Eli Lilly were purchased from over the counter.

Sample preparation

~2 ml of sample was taken out from the vial by syringe and separated into two fractions (soluble and solid fraction) by centrifuging at 500g for 15-20 minutes. The solid fraction was transferred to a 1.9 mm ZrO_2 rotor by centrifugation with the same speed using a bench top micro-centrifuge device without further altering its native state. ~4.5 mg of total crystalline protein was transferred to the rotor. ~50% water content compared to the native formulated state was present inside the rotor. Epoxy glue was used to seal the top and bottom caps in order to maintain the hydration level and other volatile excipients content. ~500 µL of supernatant (soluble fraction) from the biphasic suspension was collected and transferred to a 5 mm NMR tube (Bruker) with 50 µL D_2O for field locking without perturbing the formulation condition of the sample. The protein concentration of the soluble fraction was ~1.05 mg/ml and ~1.75 mg/ml for 70/30 and 50/50 biphasic drugs, respectively. To conduct a comparison of monomeric insulin's Higher Order Structure (HOS), the formulation was acidified. For which, 4 µL of 6N HCl was added to 1ml of Drug formulation.

NMR Spectroscopy

Both solid state and solution state NMR experiments were recorded at 298 K (sample temperature) by using Bruker 750 MHz Avance III HD spectrometer connected with variable temperature unit and internal sample temperature were calibrated with Sodium trimethylsilylpropanesulfonate (DSS). For data acquisition and processing Topspin 3.6.2 pl6

and Topspin 4.1.3 Bruker NMR software were used respectively. All solution state NMR experiments were recorded with triple resonance (TXI) room temperature (RT) probe. One-dimensional (1D) ^{1}H and ^{1}H-^{13}C Two dimensional SOFAST Methyl HMQC (2D) spectra were recorded for the R formulation as well as the soluble fraction collected from the biphasic insulin formulation. 1D ^{1}H NMR spectra were acquired using a standard Bruker pulse program ('zgesgp') with a recycling delay of 1s. As the sample concentrations were very low, 4k scans were collected with an experimental time of ~3 hours. All data were processed with exponential multiplication window function with 3 Hz line broadening, whereas States-TPPI (time proportional phase incrimination) method was used to acquire the SOFAST Methyl HMQC (2D) spectra. For R formulation data set was acquired with the size of FID 1024 × 96 points in the direct and indirect dimensions respectively and was zero filled to yield a data matrix of size 4096 × 512 before Fourier transformation. Unlike R formulation the soluble fraction of drug formulation (Insugen 50/50) was acquired using Non-Uniform Sampling (with 25% of sparse sampling) with FID size of 1024 × 256 and zero filled with matrix of 2048× 1024 due to less sample concentration present in the soluble fraction.

Solid-state NMR experiments were carried out on a 1.9 mm HCN triple resonance probe, and for better sensitivity, it was configured to double resonance mode. 1D ^{13}C Cross polarization magic angle spinning (CP-MAS) [58] experiment was recorded with an optimized cross-polarization (CP) contact time of ~1.25 ms and a ramp of 100% to 70% on the spin-lock RF power. A typical ^{1}H 90° pulse radio frequency (RF) of power 137 kHz was used. During ^{13}C acquisition of up to 21.7 ms, the ^{1}H RF power of the phase-inverted supercycled sequence for attenuation of rotary resonance (PISSARRO) [59] decoupling scheme was set to 137 kHz. The recycle delay was kept for 3 s with an experimental time of ~2.5 hrs with a number of scans of 3072. Sample processing was done using the window function (exponential multiplication) with a line broadening of 100 Hz. The line broadening can be reduced to 50-80 Hz depending on the spectra with a higher signal-to-noise ratio. However, noise may interfere during the PCA analysis if lesser line broadening is kept for those cases where S/N is not appreciable. Therefore, a line broadening of 100 Hz was kept to maintain uniformity. For the CP build-up experiment series of 1D ^{13}C CP-MAS NMR spectra with variable cross polarization contact time ranging from 10-1400 µs was recorded for each representative batch of Insugen® 70/30 and Huminsulin® 70/30. The Intensity Integrals of the aliphatic region (10-70 ppm) of the 1D ^{13}C CP-MAS NMR spectra of both I 70/30 and H 70/30 were normalized and the spectra were plotted against the CP contact time to obtain the hC-CP build-up curve.

2D ^{1}H-^{13}C HETCOR was recorded with cross-polarization (CP) contact time of ~1.25 ms and ~0.3 ms for ^{1}H to ^{13}C and ^{13}C to ^{1}H cross polarization steps, respectively with a ramp of 70% to 100% on the spin-lock RF power at a MAS frequency of 33 kHz. We used ~ 80-60 kHz of ^{1}H and 60-40 kHz ramp power to establish Hartmann-Hahn matching conditions and proton decoupling power of ~ 100 kHz during acquisition with a recycle delay of 1.5 sec. The number of scans used was 480, with 512 and 256 data points in direct (^{1}H) and indirect (^{13}C) dimensions, respectively, with an acquisition time of 4.91 ms for an experimental time of ~31hr.

Pearson's Correlation Coefficient calculation

Pearson's correlation coefficient (r) analysis provides the extent of similarity in the higher order structure between Insugen R and soluble fraction of Insugen 50/50. The correlation coefficient was calculated from the chemical shift data from the 2D ^{1}H-^{13}C SOFAST Methyl-HMQC experiment. The Pearson's Correlation Coefficient formula (shown below) was used to calculate the correlation coefficient between the drug formulations. (Note: 24 peaks out of total 28 peaks were taken for the calculation, as those four peaks were below noise level due to low concentration of the Insugen 50/50.)

$$r = \frac{\sum(x-\bar{x})(y-\bar{y})}{\sqrt{\sum(x-\bar{x})^2 \sum(y-\bar{y})^2}}$$

Combine chemical shift deviation (CCSD) calculation

The CCSD (combined chemical shift deviation) method is a targeted analysis based on the comparison of two spectra in terms of chemical shift and their intensities. The mean values of ^{1}H and ^{13}C positions (measured from the 2D ^{1}H-^{13}C SOFAST Methyl-HMQC experiments) and the variance of these means for each samples were calculated. These values were used to calculate the distance between samples using the following equation [60].

$$CCSD = \sqrt{0.5 * \left[\left(\delta_H - \delta_{H,ref}\right)^2 + \left(\alpha\delta_X - \alpha\delta_{X,ref}\right)^2\right]}$$

where δ_H and δ_X are the ^{1}H and ^{13}C chemical shifts, respectively, of a given cross peak of soluble fraction of Insugen 50/50 and $\delta_{H,ref}$ and $\delta_{X,ref}$ are the ^{1}H and ^{13}C chemical shifts of Insugen R respectively, for the same cross peak; and the frequency weighting factor, α, is equal to 0.251 for ^{13}C corresponding to the relative gyromagnetic ratio.

Principal Component Analysis (PCA)

Principal Component Analysis was performed separately for both liquid (1D ^{1}H) and solid (^{13}C CP-MAS) state NMR spectra by using Mestronova 14.2.3 software. Spectra of Four different batches of each drug class were grouped into separate classes. For the 1D ^{1}H spectra, corresponding resonance peaks for the excipients, buffer components, and water (1.044-1.137, 1.80-7.37, 9.13-10.00 ppm) were excluded from the PCA calculations. All the regions taken for PCA calculations were binned with 0.02 ppm with summed spectral intensities in each bin. The summed intensities were processed with integrity check, filtered with Inter quartile range (25%), normalized, and Pareta-scaled before PCA.

Similarly, for ^{13}C CP-MAS NMR spectra resonance peaks corresponding to aliphatic, aromatic, and carbonyl regions were considered for PCA calculation, remaining regions (0.00-10.3, 73.5-111.0, 140-153.4, 161.5-168.3, 181.0- 200.0 ppm) were excluded. All the regions taken for PCA calculations were binned with 0.4 ppm resolution with summed spectral intensities in each bin. The summed intensities were processed with integrity check, filtered with Inter quartile range (25%), normalized, and Pareta-scaled before PCA calculation.

Mahalanobis distance calculation

Mahalanobis Distance was calculated using PC scores from the PCA analysis as discussed by Wang et al.[61]. A sample matrix was made for each drug of interest for the comparison, where rows correspond to the total number of batches and columns denote the total number of Principal Components (PC). We have taken four different batches of each drug for our calculation. Three PC components were taken for 1D ^{1}H spectra and five components were considered for ^{13}C CP-MAS spectra. The covariance matrix and the Mean vector for each sample matrix were calculated. The following formula calculated Mahalanobis Distance.

$$D_M = \sqrt{(\overline{A} - \overline{B})S^{-1}(\overline{A} - \overline{B})'}$$

Where D_M denotes Mahalanobis distance, $\overline{A}$ and $\overline{B}$ denote the mean vector of the sample matrix (A and B), and S denotes the average covariance matrix of both samples matrix. $(\overline{A} - \overline{B})'$ denotes transpose matrix of $(\overline{A} - \overline{B})$.

Crystal morphology study by Bright field microscopy

2 µl of the drug was placed in a glass slide in its native formulation. Coverslips were mounted above the sample, followed by immersion oil. The crystal images were captured under a laser scanning microscope (Carl Zeiss, LSM 780) with 60X zoom. We measured the length of 100 crystals from each drug and plotted it in a Gaussian distribution using Image J and OriginPro (2021) software, respectively.

Wide angle X-ray Scattering (WAXS)

WAXS data for Insugen® 70/30 and Huminsulin® 70/30 were collected at Xenocs 2.0 SAS instrument using an Eiger R1M vacuum fed through a set high-resolution hybrid pixel photon-counting detector. Samples were transferred to a capillary tube and kept at 4ºC overnight to settle the crystals. The X-ray scattering pattern was collected from the bottom of the capillary after exposure for two hours at room temperature. The distance between sample-to-detector was 333.307 mm, and the range of scattering angle (2θ) 1.5-5º was covered. The buffer scattering was subtracted from the original crystal scattering. OriginPro (2021) software was used for data processing and plotting.

Invitro release of Insulin from the crystal

The release of Insulin hexamer from crystals was observed for 13 hours. 2 ml of each sample was centrifuged, and crystals were collected after the supernatant was discarded. Crystals were dissolved in 25 ml of release buffer (50 mill molar phosphate buffer, pH 7.4 with 0.01% tween-80) in a 50 ml falcon tube at 37ºC and shaken at 50 rpm in an incubator. 25 µl of aliquot were taken out in triplicate, and mixed with 200 µl of working BCA reagent, and absorbance of the mixture (copper-BCA complex) was recorded at 562 nm after 3 and 13 hours. The Insulin standard curve was made by plotting the known concentrations of Insulin against its absorbance at 562 nm. All the experiments were carried out in triplicate, and a statistical t-test

Chemical Shift Predictions

The assignment of 2D ^{13}C-^{1}H HETCOR spectra was done from predicted chemical shifts of insulin hexamers using SHIFTX2 program version 1.10 (http://www.shiftx2.ca/index.html).

PDB ID: 1ZNJ (R6) and 1MSO (T6). A pH value of 7 and temperature of 298 K were used for the predictions.

Results

Human Insulin in the biphasic microcrystalline suspension (dispersed microcrystalline solid in a liquid) drug exists in the hexameric state in the presence of metal ion Zinc and m-cresol, which has been well established through X-ray crystallography [62]. While the hexameric state enhances the stability of Insulin, the microcrystalline suspension formed due to protamine sulfate leads to the protracted release of Insulin. Nonetheless, the monomeric state of Insulin is biologically active and participates in its functional activity ***(Figure 5A)***.

HOS characterization of biphasic microcrystalline in the formulated state is challenging as the formulation exists in soluble and insoluble fractions. Also, the solid fraction (microcrystalline protein) settles down, leaving the soluble fraction as a supernatant when the formulation is kept for 10-15 minutes. We exploited this property and made the process easier; we separated both fractions by mild centrifugation. The soluble formulation of Insulin is known to preserve the hexameric state of the Insulin in the presence of zinc and m-cresol. The 1D ^{1}H spectra of the soluble fraction of Huminsulin® 70/30 were compared with the intact biphasic suspension without any perturbation to investigate any possible impact on higher-order protein structure during the separation of the two components. We found no differences in the spectral distribution in methyl and amide regions indicating, the HOS in the soluble fraction is well preserved after the separation by centrifugation ***(Figure 1)***.

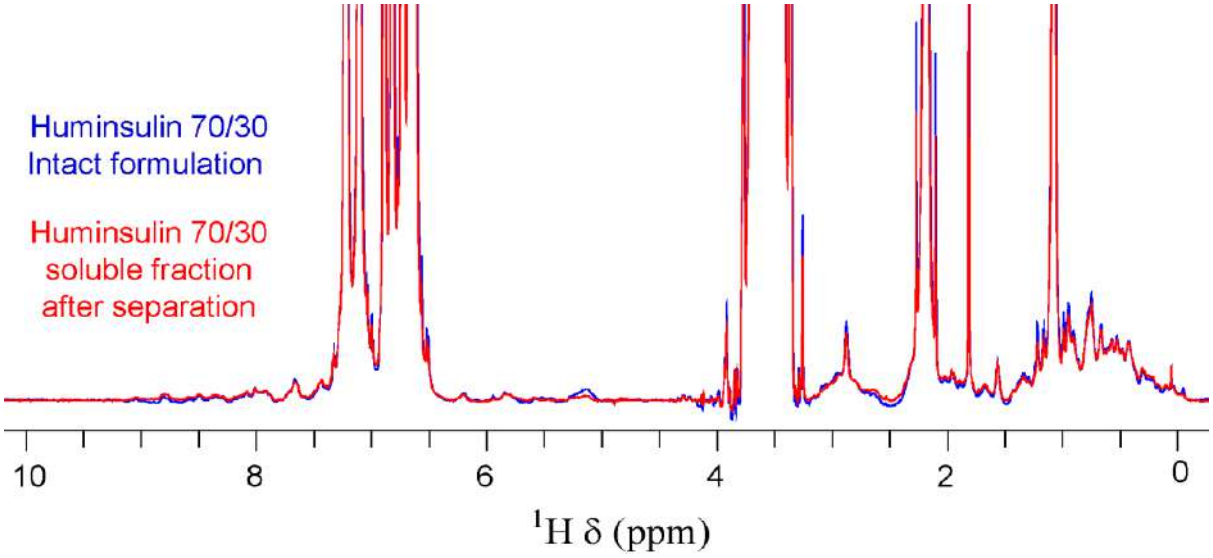

Figure 1. Investigation of HOS of soluble fraction after centrifugation: *Higher order structure similarity of Huminsulin®70/30 biphasic suspension intact (blue) (data were recorded on samples as it is, without any separation of suspension part) and soluble fraction of Huminsulin 70/30 after centrifugation at 500 g (red).*

We performed comparative study of 2D ^{1}H-^{13}C SOFAST Methyl HMQC NMR experiments on the soluble fraction of Insugen 50/50, Insugen 70/30 and compared with in Insugen R

(Regular formulation) to further validate the stability of HOS of the soluble fraction after separating from solid fraction by centrifugation.

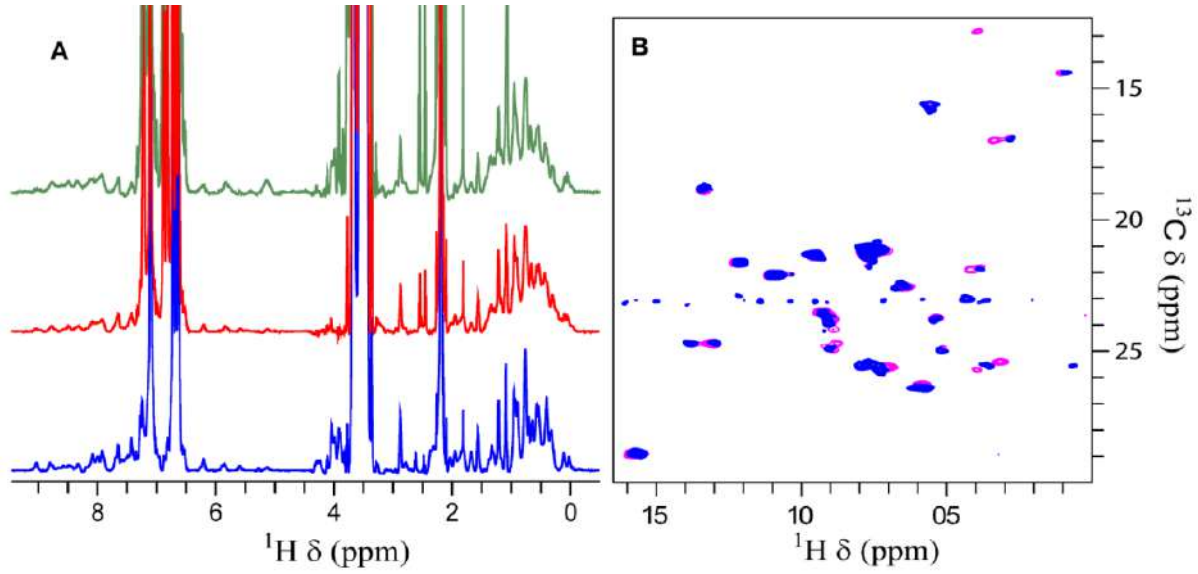

Figure 2. Comparison of HOS of soluble fraction to regular formulation: *Higher order structure similarity of soluble fraction of the biphasic suspension to the Regular formulation (R) shown by 1D ^{1}H and 2D ^{1}H-^{13}C SOFAST Methyl HMQC NMR spectroscopy. (A) 1D ^{1}H NMR spectra of Insugen R (blue), soluble fraction of Insugen 50/50 (red) (2-fold increased spectral intensity), soluble fraction of Insugen 70/30 (green) (2-fold increased spectral intensity). (B) Overlapped 2D ^{1}H-^{13}C SOFAST Methyl HMQC NMR spectra of Insugen R (pink) and soluble fraction of Insugen 50/50 (blue). [Note: as the concentration of Insugen 50/50 was less compared to Insugen R, 2D ^{1}H-^{13}C SOFAST Methyl HMQC NMR was carried out with 25% Non-Uniform Sampling Scheme (NUS)]*

The peak position in the aliphatic (1.044-1.137 and 1.80-7.37 ppm) and amide (9.13-10.00 ppm) regions are similar in all the spectra, however very minute changes are seen between the soluble fractions and the regular formulation due to differences in the excipients in the formulations ***(Figure 2A)***. Further 2D ^{1}H-^{13}C SOFAST Methyl HMQC NMR ***(Figure 2B)*** between Insugen R (pink) and soluble fraction of Insugen 50/50 (blue) shows a promising superimpose which indicates similarity of HOS. Minute differences were seen in the 2D NMR spectra as well, reflects the 1D NMR spectra due to the different excipients condition. We performed Pearson correlation coefficients (r) analysis between the proton & carbon chemical shift ***(Figure 3 A, B)*** and found 0.999 for both shows a good correlation. 4 peaks out of 28 peaks in soluble fraction of Insugen 50/50 were below noise level were excluded from the calculation. We also used Combine Chemical Shift Difference (CCSD) analysis ***(Figure 3C)*** as a statistical tool and found the CCSD values were below the cut-off 0.03 which also shows the significant HOS similarity.

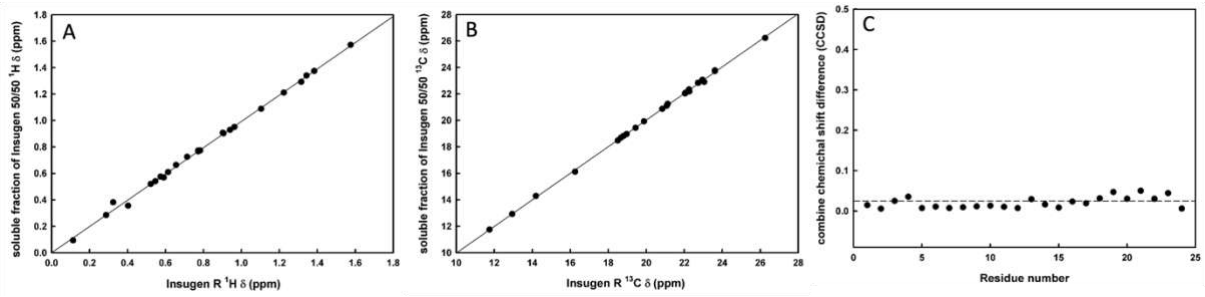

Figure 3. Statistical Comparison of HOS of soluble fraction to regular formulation: *(A). ^{1}H Correlation coefficient analysis between Insugen R and soluble fraction of Insugen 50/50. (B) ^{13}C Correlation coefficient analysis between Insugen R and soluble fraction of Insugen 50/50. Both shows a correlation of 0.999, denotes a good correlation between the drug products. (C) Combine chemical shift difference between Insugen R and soluble fraction of Insugen 50/50. Dash line define a cut-off of 0.03 ppm, where CCSD of maximum residue falls below the cut-off shows a great correlation with minute differences in few residues*

From all above experiments and analysis, we were confirmed that the high order structure of soluble fraction was preserved after the centrifugal separation. Further, we assess the water content of the crystalline fraction inside the rotor as the hydration level plays a crucial role in maintaining the crystalline structure. We Performed 1D ^{13}C CP-MAS for native NPH formulation (100% crystalline) without centrifugation and compared the spectral quality in different hydration levels for Huminsulin ***(Figure 4)***.

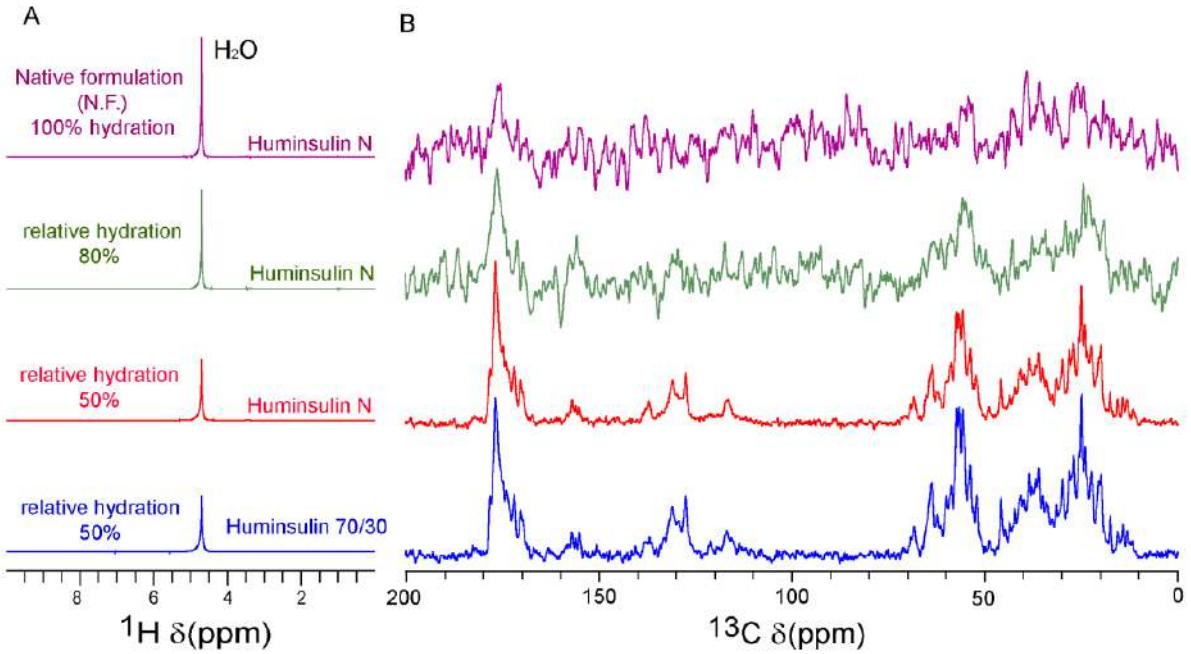

Figure 4. HOS of crystalline insulin under different hydration level:*(A) ^{1}H proton-detection shows water content present in the rotor. The native formulation is considered as 100 % hydration (magenta), 80 % relative hydration compared to the native formulation (green), 50% relative hydration (red), and 50% relative hydration compared to the native formulation for Huminsulin (blue), (B) (magenta) 1D ^{13}C CP-MAS NMR spectra of the Huminsulin® N without centrifugation. The sample concentration was very low (~40 µg); therefore 30 K scan was recorded. (green) 1D spectra of the Huminsulin® N concentrated ~10 times (~400 µg) with*

80% water content compared to the native formulation; here, 10 K scan was recorded, and 2.09 of S/N ratio was observed. (red) 1D ^{13}C CP-MAS spectra of the Huminsulin® N with 50% water content on ~5 mg of sample was recorded for 3k scans, where S/N ratio of 13.45 was observed. (blue)1D ^{13}C NMR spectra of the Huminsulin® 70/30 with 50% water content with for 3k scans, here. ~4.5 mg of sample was packed inside the rotor and S/N ratio of 12.77 was observed. Thus, the 50% water content of the native formulation, is well suited to maintain the higher- order structure of insulin inside the rotor.

For native NPH formulation, the water content was considered to be 100%, and the protein content inside the rotor was ~40 µg. We recorded a 30 K scan and can achieve an S/N ratio of only 1.5. However, we believe it may not be a commercially viable option to perform such an analysis in the native state at natural abundance due to low S/N. We further concentrated the crystalline part by mild centrifugation and made it to 80% and 50% relative hydration levels compared to the native formulation inside the rotor Clearly, a better S/N was observed for these samples. At 50% relative hydration, we observed good S/N with a value of > 10 in a reasonable experimental time of ~2.5 hrs, which is typically required during the quality assessment of pharmaceuticals. Except for better resolution and sensitivity, spectra were found to be comparable at different hydration levels, indicating that our choice of 50% hydration is optimal. It does not impede the crystalline structure of the native formulated state, therefore we carried out all our experiments with 50% of relative hydration to the native formulated state.

HOS assessment for the soluble fraction of the Insulin biphasic formulation

1D ^{1}H NMR is a high-resolution technique that provides valuable information about the overall protein structure and its folded state. Jack E *et al*. have shown the usefulness of 1D ^{1}H NMR to investigate the stability of protein and excipients in the native biopharmaceutical formulation [63]. However, 1D ^{1}H NMR is limited to biopharmaceuticals in a solution formulation. Typically, crystalline suspensions are converted to soluble states to investigate using liquid-state NMR, which might introduce structural perturbation. Whereas, in this study, because of the biphasic nature, the soluble and solid fractions of the biphasic microcrystalline formulation were separated by mild centrifugation without perturbing the native state ***(Figure 5B)*** and were assessed independently by liquid and solid-state NMR, respectively.

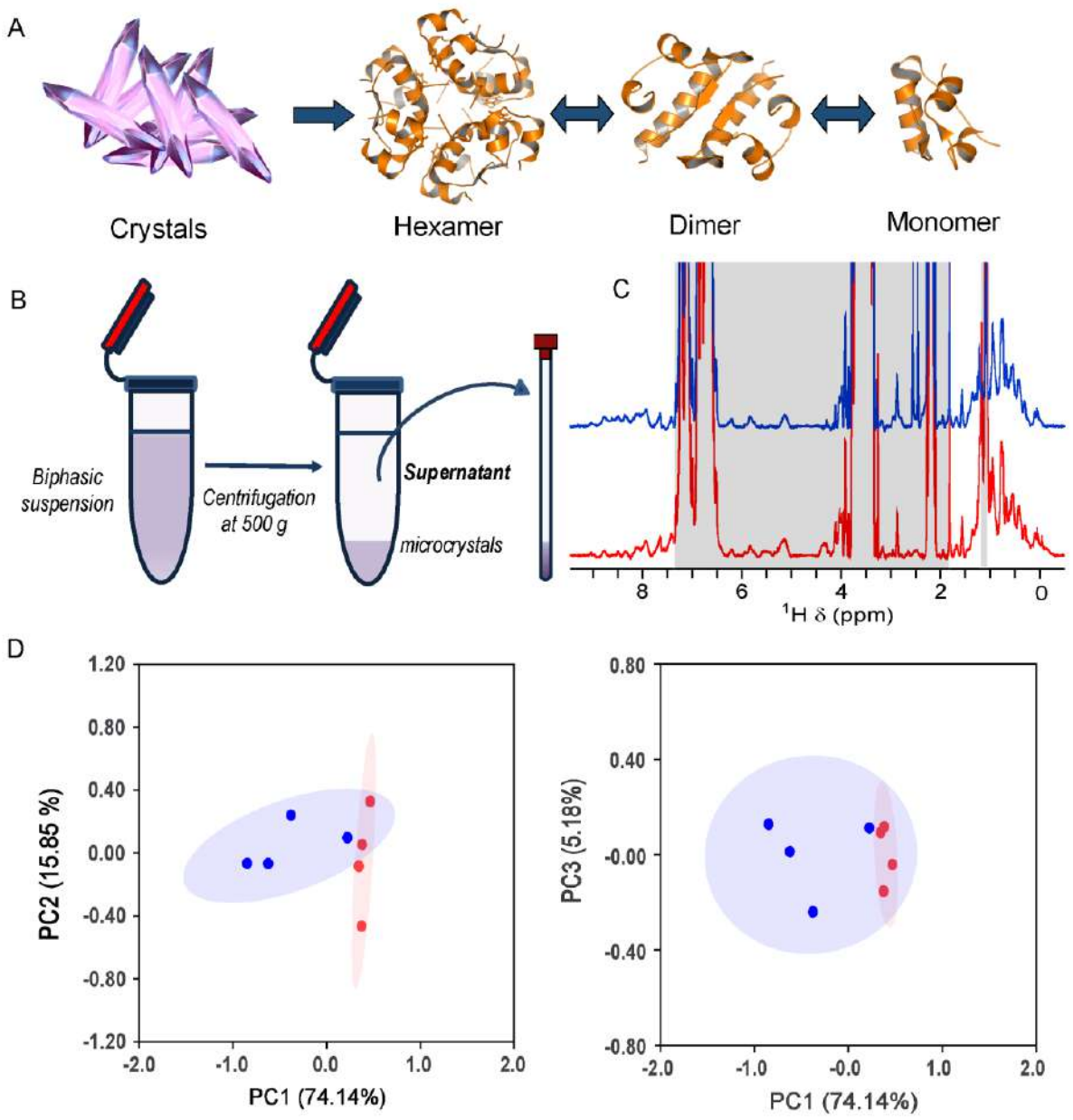

Figure 5. HOS assessment for the soluble fraction of the Insulin 70/30 formulation:*(A) Cartoon mimicking Insulin hexamer releases from the crystal, which further converts to dimer and monomer (active form). (B) Schematic of Higher order structural similarity comparison of soluble fractions of the biphasic suspension by 1D ^{1}H NMR spectroscopy. (C) 1D ^{1}H NMR spectra of soluble fraction of Huminsulin® 70/30 (red) stacked with soluble fraction of Insugen® 70/30 (blue). PCA calculation was carried out for aliphatic and amide regions. The excipients (shaded in grey) spectra were excluded from the calculation. (D) Score plot of principal component 1 (PC1) vs. principal component 2 (PC2) and principal component 1 (PC1) vs. principal component 3 (PC3) of the soluble fraction of Huminsulin® 70/30 (orange) and soluble fraction of Insugen® 70/30 (blue). Blue and red ellipses are drawn with 95% CI for Insugen® and Huminsulin®, respectively.*

1D ^{1}H NMR on the soluble fraction (SF) of Huminsulin® 70/30 (H 70/30) and Insugen® 70/30 (I 70/30) was performed ***(Figure 5 C)***. A good spectral dispersion (~ 3 ppm) was observed in

the amide region, indicative of a well-folded protein. The spectral fingerprint in the methyl and amide regions was comparable between the two drug products. Though qualitative comparison was informative, we performed Principal component analysis (PCA) for quantitative assessment. PCA is a well-established chemometric approach often used for comparability assessment of biotherapeutics [61]. For PCA calculations, we recorded 1D ^{1}H NMR for four different lots of each drug product. To overcome any impact of firm excipients peaks, spectral regions only specific to human Insulin peaks were considered for the PCA analysis. The first three PC components accounted for > 90% of spectral variations. PCA score plots between SF of H 70/30 and SF of I 70/30 ***(Figure 5D)*** showed a promising cluster similarity. Using these PC scores, Mahalanobis Distance (D_M) was calculated, and it was found to be 2.69. Considering D_M of 3.3 as an acceptance similarity metric criterion; for many regulatory agencies (e.g., FDA, EMA), also shown by Wang et al. [61], both drug products showed a comparable higher order structure between Insugen® 70/30 and Huminsulin® 70/30 in the soluble fractions. A similar study was performed with the Insugen® 50/50 and Huminsulin® 50/50 ***(Figure 6)***, and the results were similar to the above.

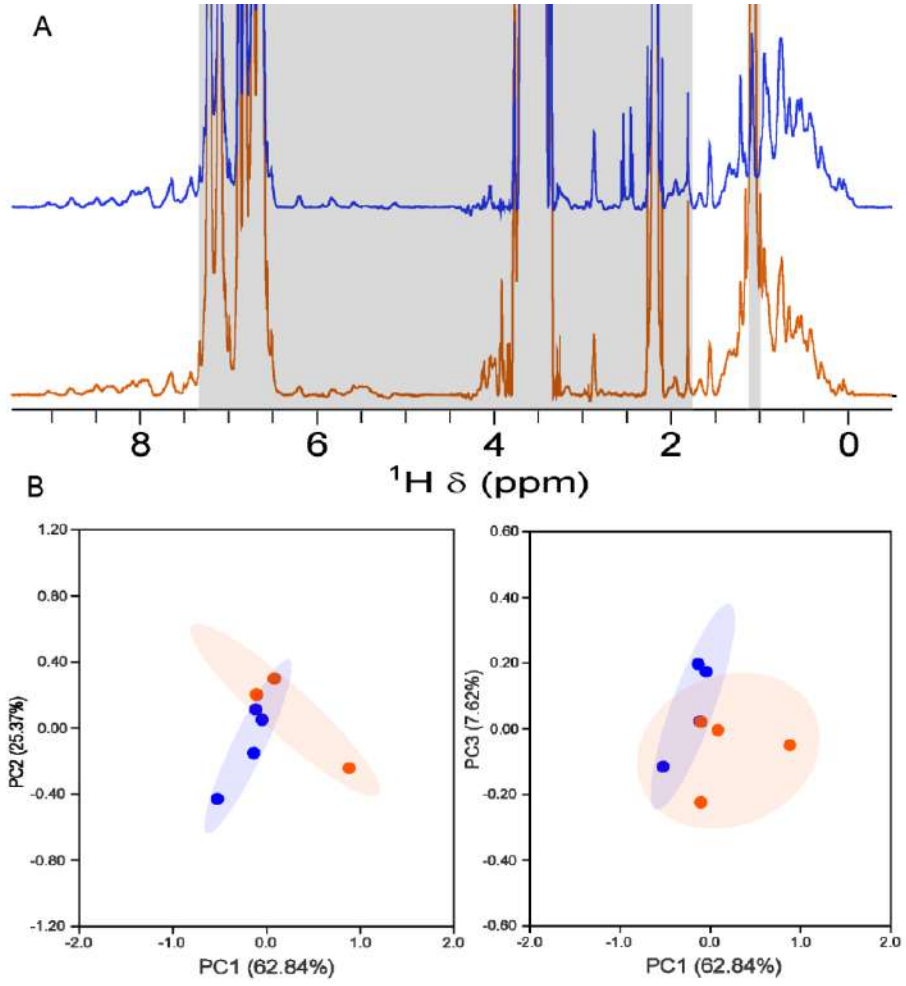

Figure 6. HOS assessment for the soluble fraction of the Insulin 50/50 formulation*: (A) 1D ^{1}H NMR spectra of the soluble fraction of Huminsulin®50/50 (orange) stacked with a soluble fraction of Insugen®50/50 (blue), (B) Score plot of principal component 1 (PC1) vs. principal component 2 (PC2) and principal component 1 (PC1) vs. principal component 3 (PC3) of the soluble fraction of Huminsulin®50/05 (orange) and soluble fraction of Insugen®50/50 (blue).*

HOS assessment for the suspension fraction of the Insulin biphasic formulation

1D ^{13}C CP-MAS NMR is a powerful non-invasive, non-destructive structural characterization tool for samples observed at natural abundance, most beneficial for biopharmaceutical sample [64], where isotope labelling of samples is challenging. A previous study by Mingyue *et al.* investigated the crystallinity and crystal stability of microcrystalline pembrolizumab upon dehydration and change in temperature by using MAS NMR [64]. We carried out all solid experiments with ~50% hydration level of the native formulation for a better S/N ratio and minimum possible experimental time.

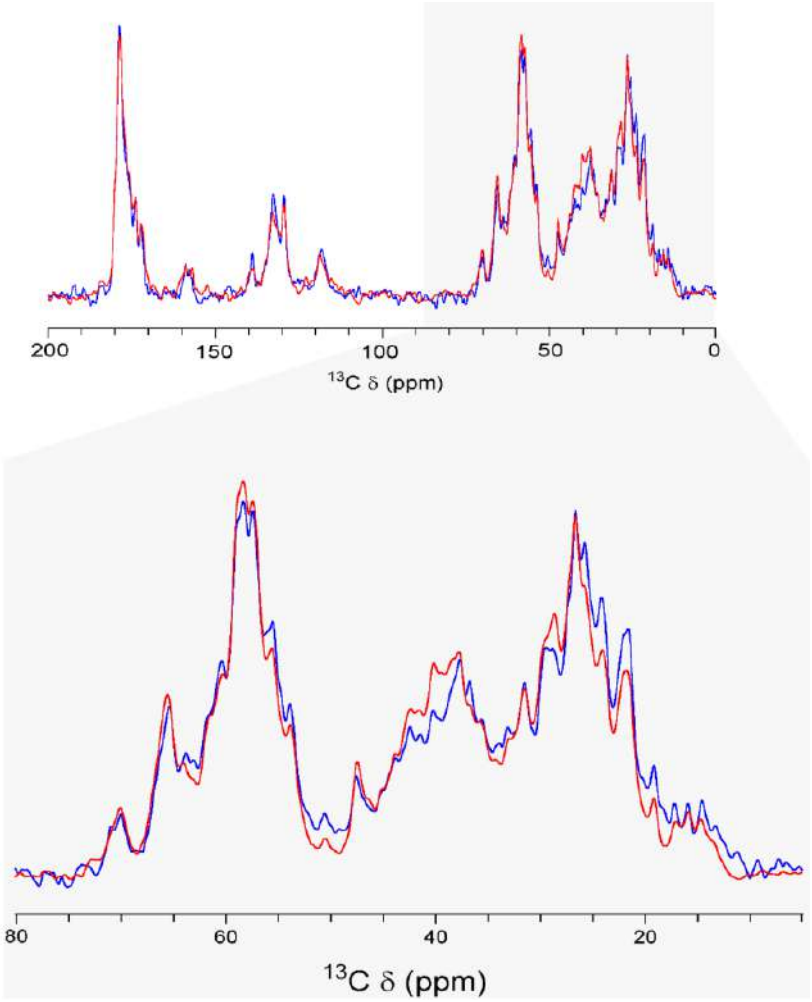

Figure 7. 1D ^{13}C CP MAS spectra comparison of Huminsulin® 70/30 at different spinning speeds*; at 33 kHz spinning speed (blue), and at 20 kHz spinning speed (red). The top image shows the complete ^{13}C spectra, while the bottom image shows the zoomed region from aliphatic carbons.*

The initial set of experiments was carried out to assess the suitability of lower MAS rates for more commonly available MAS rotors for routine assessment in drug quality control. We conducted a comparative study of 1D ^{13}C CP MAS experiments at different spinning MAS rates of 20 kHz and 33 kHz. At 33 kHz, the spectral peaks were a little sharper compared to 20 kHz; however, we did not observe significant differences in the linewidth of the spectra at both spinning speeds ***(Figure 7).***

Considering the higher availability of lower MAS rates rotors in drug quality control, we suggest a spinning speed of 20 kHz suffices to investigate the molecular level structural details and dynamics of a suspension formulation of molecules ranging from small proteins up to 40 kDa. (Insulin hexamer ~36 kDa). As discussed above, for the suspension portion containing microcrystalline particles ***(Figure 8A),*** we performed 1D ^{13}C cross-polarization at the MAS spinning rate of 20 kHz for structural assessment.

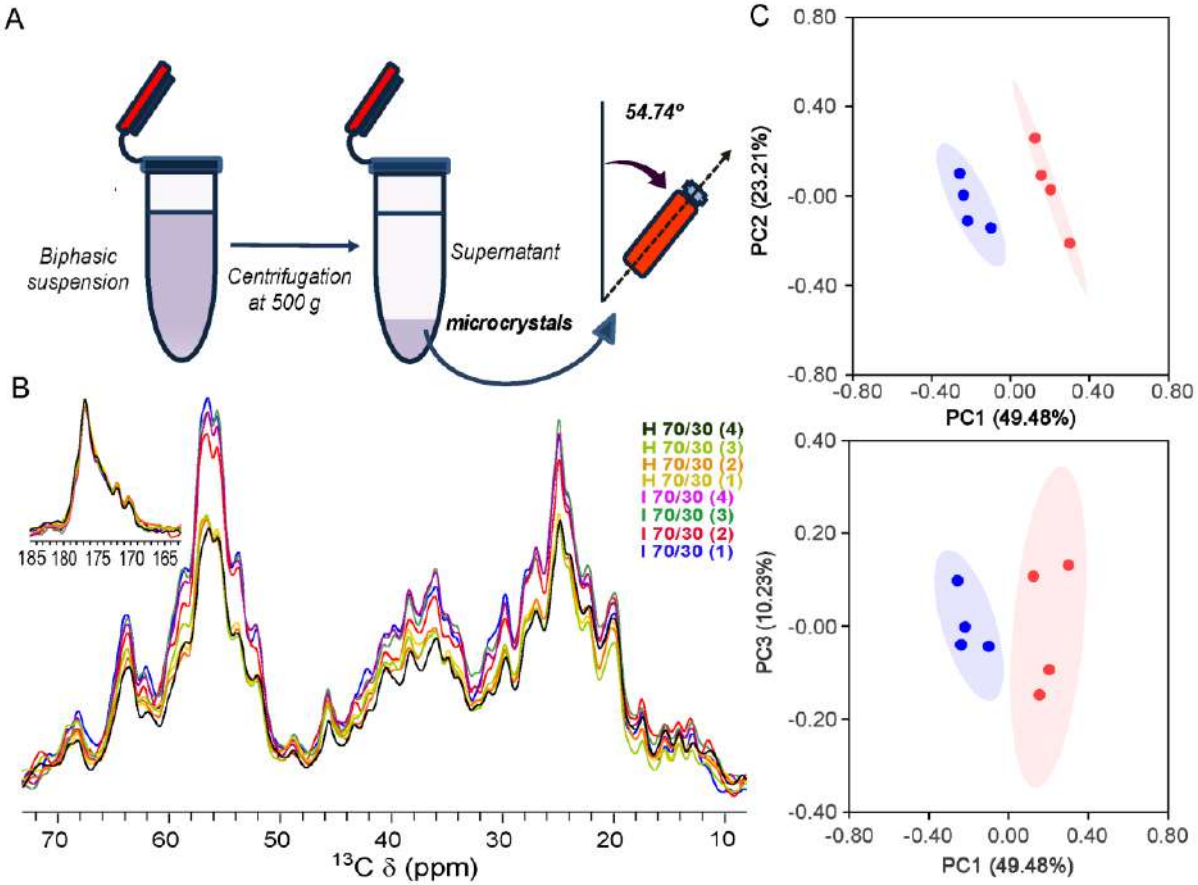

Figure 8. HOS assessment for the suspension fraction of the Humulin70/30 vs Insugen 70/30 formulation: *(A) Schematic of Higher order structural similarity comparison of solid fractions of the biphasic suspension by 1D ^{13}C CP-MAS NMR. (B) All four 1D ^{13}C CP-MAS NMR spectra of the solid fraction of Insugen® 70/30 (blue, red, green, magenta) superimposed with a solid fraction of Huminsulin® 70/30 (yellow, orange, light green, dark green). The Carbonyl region of all the spectra was normalized, as shown in the small inset. (C) Score plot of principal component 1 (PC1) vs. principal component 2 (PC2) and principal component 1 (PC1) vs. principal component 3 (PC3) of the solid fraction of Insugen® 70/30 (blue) and solid fraction of Huminsulin® 70/30 (orange). Blue and red ellipses are drawn with 95% CI for Insugen® and Huminsulin®, respectively. Full-length spectra are provided in the supporting material.*

The qualitative assessment of overlaid spectra of 1D ^{13}C CP MAS showed a high degree of fingerprint match between both drug products, suggesting the HOS structure of Insulin inside the crystals is highly similar. However, in the aliphatic region of the spectra, we could observe an intensity difference (when the carbonyl regions were normalized) ***(Figure 8B).*** Further, we

recorded four different batches of each drug product. As specified above, regions specific to aliphatic, aromatic, and carbonyl peaks were selected for PCA calculations, followed by Mahalanobis distance calculation [61].

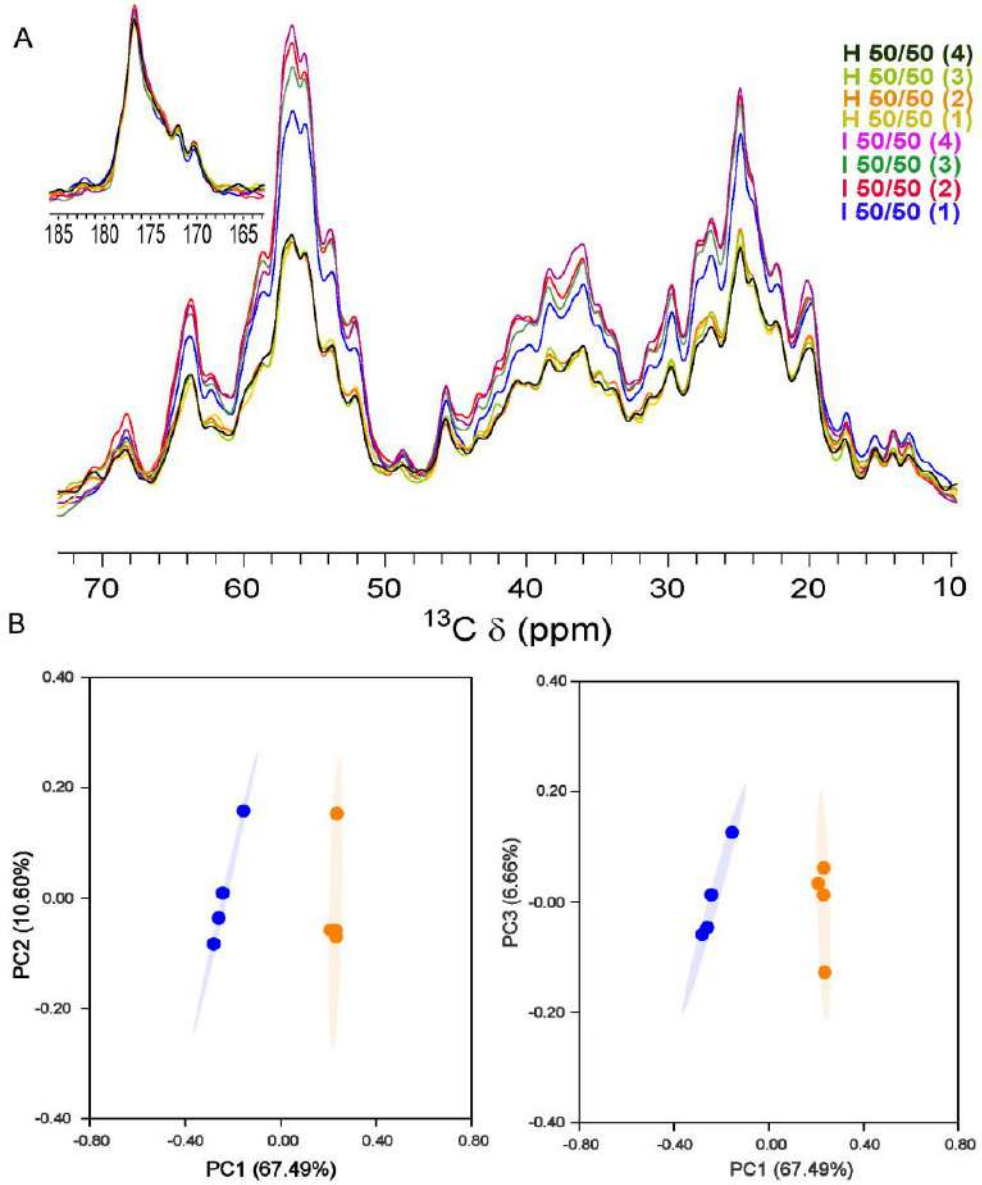

Figure 9. HOS assessment for the suspension fraction of the Humulin50/50 vs Insugen 50/50 formulation: *(A) 1D ^{13}C CP-MAS NMR spectra of the solid fraction of Insugen®50/50 (blue, red, green, magenta) different batches superimposed with a solid fraction of Huminsulin®50/50 (yellow, orange, light green, dark green) four different batches. The Carbonyl region of all the spectra was normalized. (B) Score plot of principal component 1 (PC1) vs. principal component 2 (PC2) and principal component 1 (PC1) vs. principal component 3 (PC3) of the solid fraction of Insugen®50/50 (blue) and solid fraction of Huminsulin®50/50 (orange) [carbonyl region of all the 1D ^{13}C CP-MAS NMR spectra were normalized].*

Unlike the soluble fraction, the PCA plots did not show any overlap (Figure 8C), and the Mahalanobis distance (D_M) for the suspension fraction was found to be 40.83, clearly showing

the differences in the two drug products. A similar study on the 50/50 biphasic drug yielded a D_M value of 117.22, indicating differences in the two sets of samples ***(Figure 9).***

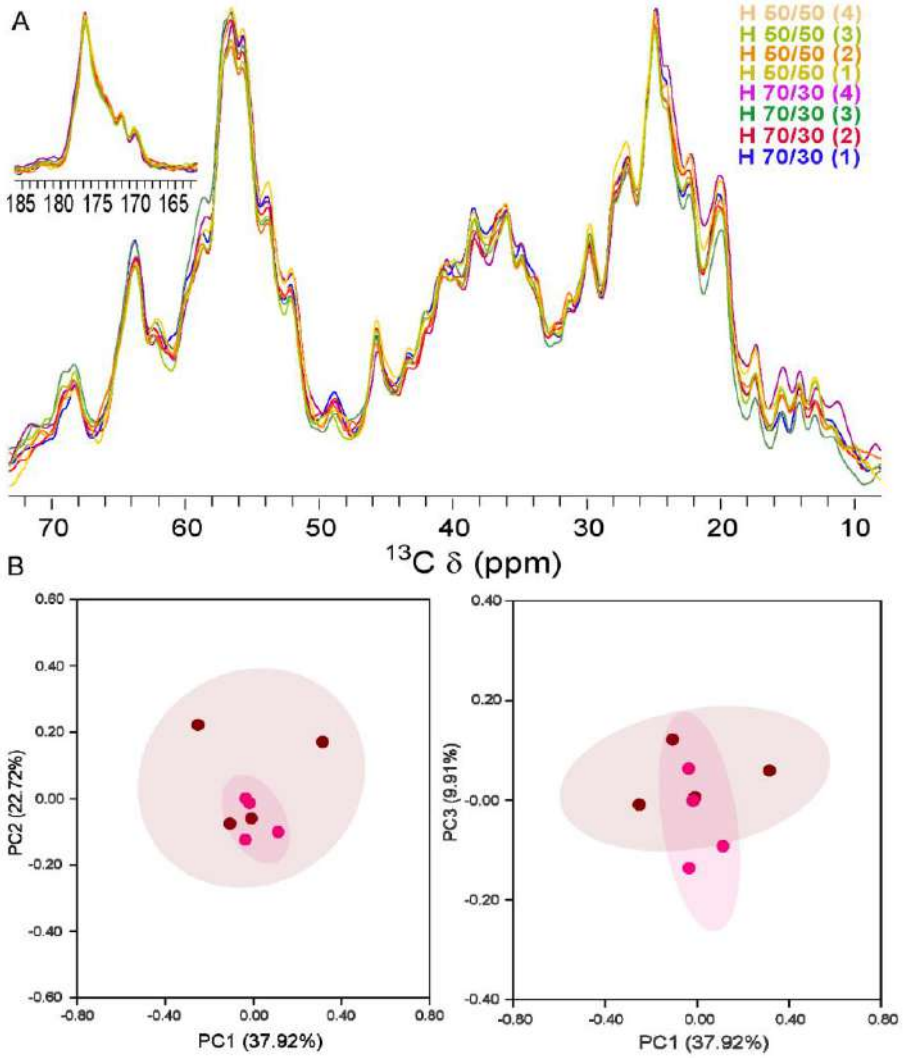

Figure 10. HOS assessment for the suspension fraction of the Humulin 70/30 vs Humulin 50/50 formulation: *(A) 1D ^{13}C CP-MAS NMR spectra of the solid fraction of Huminsulin®70/30 (blue, red, green, magenta) from different batches superimposed with a solid fraction of Huminsulin®50/50 (yellow, orange, light green, dark green) from different batches. The Carbonyl region of all the spectra was normalized. (B) Score plot of principal component 1 (PC1) vs. principal component 2 (PC2) and principal component 1 (PC1) vs. principal component 3 (PC3) of the solid fraction of Huminsulin®70/30 (brown) and solid fraction of Huminsulin®50/50 (pink). [carbonyl region of all the 1D ^{13}C CP-MAS NMR spectra were normalized.*

To verify the difference in the PCA and high D_M values only because of the difference in intensity in the aliphatic region, we replicated the study in the crystalline protein of the same brand (as control) between Huminsulin® 70/30 and Huminsulin® 50/50 and interestingly observed no differences in the spectral pattern as well as the intensity in the aliphatic region.

The PC plots were found well overlapped with each other, and the D_M value was found to be 2.52, suggesting there are no differences in the drug products ***(Figure 10).***

The intensity of the aliphatic region of Huminsulin® 70/30 was less compared to Insugen® 70/30. As the cross-polarization (CP) dipolar coupling transfer is directly proportional to the molecule's rigidity, it indicates the aliphatic region of Huminsulin® batches is more flexible than the Insugen®, resulting in less cross-polarization transfer. To further validate this observation, we recorded a series of 1D ^{13}C CP MAS experiments with variable CP contact time (ranging from 10-1400 µs) to obtain the hC-CP build curve. The Intensity integrals of the aliphatic region (10-70 ppm) were normalized and plotted against the CP contact time for both Insugen® 70/30 and Huminsulin® 70/30. We observed a marginally faster build-up curve with oscillation for I 70/30. The CP build-up reaches the maximum intensity earlier in I 70/30 (400 µs) compared to H 70/30 (600 µs) ***(Figure 11A, B).*** It is well established that if there is dominant dipolar interaction, the polarization transfer shows oscillatory behaviour [65,66]. The frequency of transient oscillation is directly proportional to the dipolar coupling constant, which reflects the mobility or rigidity of nuclei involved therein [67]. A higher transient oscillation is directly proportional to the rigidity of the molecule43,44. Higher oscillation in the build-up curve suggested that nuclei in I 70/30 are relatively more rigid compared to H70/30. The correlation of chemical shift and heteronuclear dipolar coupling interactions has been studied using pulse sequence such as polarization inversion spin exchange at the magic angle (PISEMA) [69] to determine the peptide structure in an oriented membrane. Similarly, several studies used different modes of magnetization transfer, such as dipolar-coupling-based or J-coupling-based mechanisms for spectral editing for the rigid or flexible part of the protein [70]. To further investigate any differences in the quality of the two drug products and delineate the origin of the difference appearing in various formulations, we performed additional assessment by 2D ^{1}H-^{13}C HETCOR NMR (ssNMR) and wide-angle x-ray scattering (WAXS) on a representative batch of the two products.

2D ^{13}C-^{1}H HETCOR spectral comparison

In 1D CP MAS experiments, resolution in the carbon dimension was limited, which could be overcome by recording 2D (^{13}C,^{1}H) correlation spectra. Due to higher sensitivity, ^{1}H is the preferred nucleus of detection; therefore, we used a proton-detected ^{1}H-^{13}C 2D heteronuclear correlation (HETCOR) [71]experiment with improved resolution. Such experiments have been performed on Insulin [72] and mAbs [64] to gain structural insights into these classes of formulated biopharmaceuticals. From the 2D ^{1}H-^{13}C HETCOR spectral comparison, we did not observe a significant difference in the two spectra. However, we could even see that CP transfer around 25 ppm (methyl's) is significantly higher for ^{13}C resonances. That resulted in the appearance of more peaks and the spectral resolution in the methyl region and Cα region (50-70 ppm) in the w_1 (^{13}C) dimension for the I 70/30 ***(Figure 11 C, D)*** than H 70/30 ***(Figure 11 E, F).***

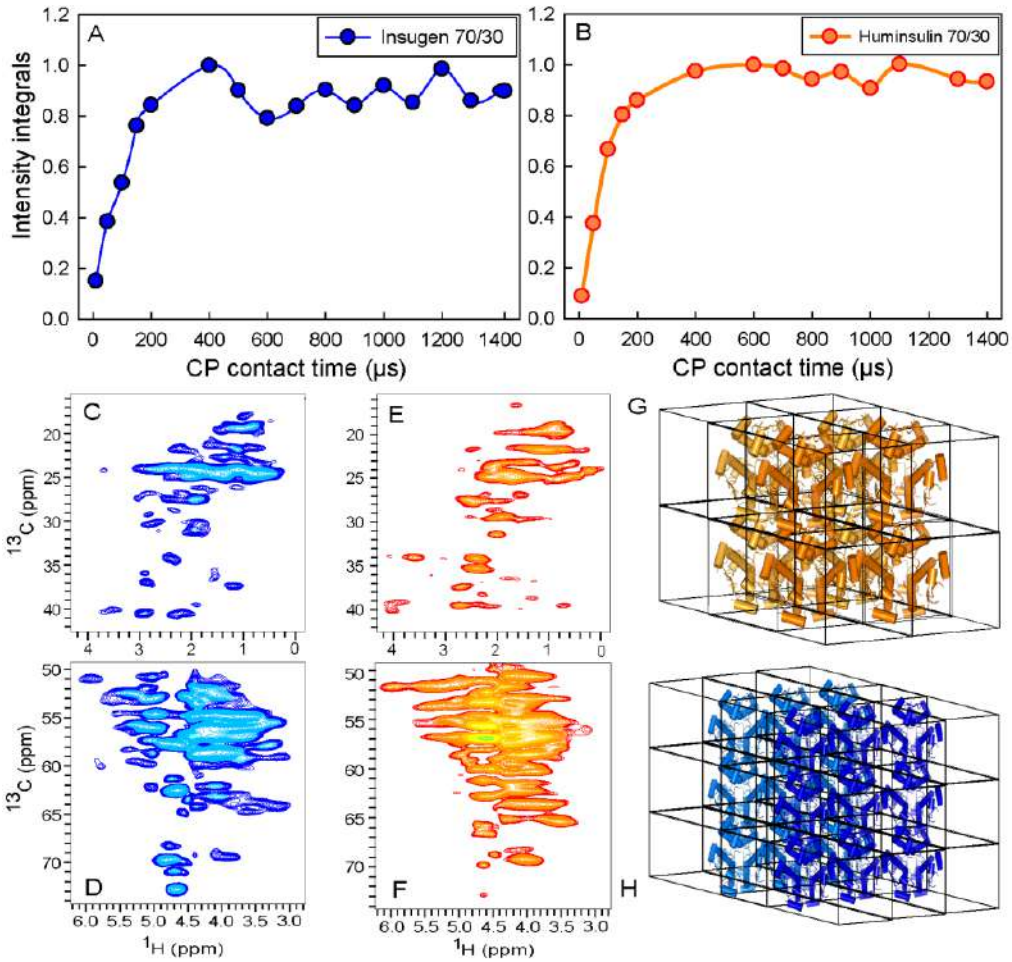

Figure 11. hC-CP build-up curve of Humulin vs Insugen: *hC-CP build-up curve of aliphatic regions (10-70 ppm) of the solid fraction of Insugen® 70/30 (A) and Huminsulin® 70/30 (B). 2D ^{1}H-^{13}C CP MAS HETCOR spectra of Insugen® 70/30 (C, D) vs. Huminsulin® 70/30 (E, F). Schematic of the packing of Insulin in the crystallite, showing more compactly packed in I 70/30 (H) and corporately loosely packed in H 70/30 (G).*

This also corroborated with the observation from the 1D ^{13}C CP MAS spectra. From all the above experiments, we observed Insulin molecules present in I 70/30 are more rigid compared to H 70/30 in the crystal lattice, allowing efficient CP transfer in the side chains of the amino acids. This could potentially be due to slightly different packing in its crystalline state. We hypothesized that the number of insulin molecules in one-unit crystal area is more in the case of I 70/30, making the crystal packing more compact than H 70/30 ***(Figure 11 G, H).*** We carried out wide-angle X-ray scattering experiments for both microcrystals to further verify our speculation on crystal packing.

Characterization of microcrystals through WAXS

X-ray crystallography is a well-established technique extensively used for investigating protein microcrystals [62]. Balschmidt *et al*. have shown the preparation of the NPH (Neutral Protamine Hagedorn) formulation of insulin hexamer containing two zinc atoms and one protamine molecule at pH 7.3 in the tetragonal crystal system (space group $p4_32_12$). However, X-ray

diffraction on a single crystal of a biopharmaceutical sample is very difficult due to the smaller size. In a study, Mathias *et al*. performed an X-ray diffraction study of different commercial and in-house Insulin microcrystalline formulations by using a synchrotron. We recorded scattering profiles from angles ranging from 1.5-5º for I 70/30 and H 70/30 microcrystals. Multiple characteristic peaks at approximately similar 2θ values were found for both samples, suggesting both the crystals are made up of unit cells of a similar space group [23] ***(Figure 12 A).*** However, further analysis of the first-order Bragg peaks (which are centered at 2θ =2.12 for both the samples) provided a full width at half maxima (FWHM) of 0.11691 and 0.08202 for I 70/30 and H 70/30, respectively ***(Figure 12 B).***

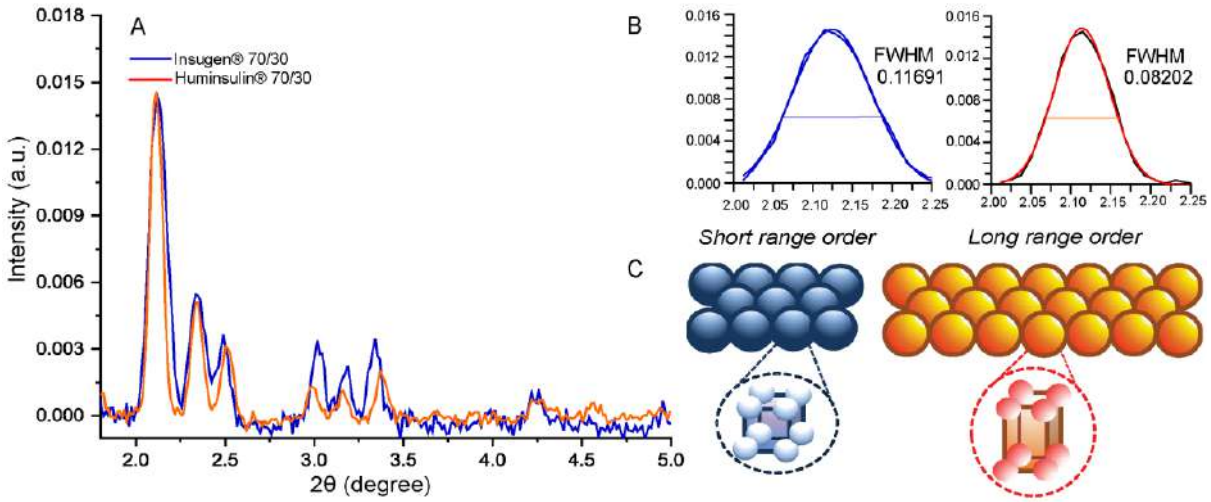

Figure 12. SAXS analysis of Humulin and Insugen microcrystal *(A) Comparison of the diffraction pattern of Huminsulin® 70/30 and Insugen® 70/30 crystals (Spectral intensities were normalized). (B) Comparison of the full width at half maxima (FWHM) at peak 2.12° of I 70/30 (blue) and H 70/30 (orange). (C) Comparison of range order of crystallite packing, which was found short-range order and long-range order for I 70/30 (blue) and H 70/30 (orange), respectively.*

This indicated a difference in the long-range order of the crystals in both samples. Now, the Debye-Scherrer analysis [73] of the Bragg peaks can be used to obtain the crystallite size (D) viz. the long-range order in the material and is given by $D= K\lambda/(\beta \cos \theta)$, where K is Scherrer's constant (K=0.94), λ is the X-ray wavelength (1.54 Å), β is full width at half maxima (FWHM) of the diffraction peak, and θ is diffraction angle. Accordingly, the crystallite size for I 70/30 and H 70/30 was 12.39 nm and 17.67 nm, respectively (see appendix for calculation). This suggested that although the crystal was of a similar space group, the H 70/30 exhibited a higher long-range order as compared to I 70/30 ***(Figure 12 C).*** Further, the FWHM of the subsequent peaks follows a similar trend, confirming the higher long-range order of H70/30. Higher order Bragg peaks (in the range 2θ = 2.4-3.5º) for I 70/30 and H 70/30 showed a slight deviation in the peak maxima, suggesting the difference in the unit cell lengths although the space group remained similar ***(Figure 12 C).*** This is in agreement with the previously reported studies on Insulin crystals [23]. So the spatial arrangement of Insulin hexamer in I 70/30 are close to each other (as the length of the unit cell is shorter) inside the unit cell of the crystal and make it more

compact compared to H 70/30. The observations stemming from WAXS and ssNMR corroborated sufficiently well and suggested a difference in the level of the crystal packing of the microcrystalline formulation.

Crystal morphology study and Invitro release of Insulin from the crystal

Microcrystalline suspension of Insulin is prepared for sustained release of the drug. The dissolution of the crystal plays a rate-limiting factor in the absorption of the drug into the blood stream [23]. Thus, careful investigation of the physical and chemical properties of the crystals is needed to maintain the homogeneity of the crystalline formulation. As we observed the difference in the crystal packing, we investigated the morphology of the crystals using bright field microscopy. We found similar crystal morphology, long and pointed-ended ***(Figure 13A, B)***, and a size distribution of 5-9 µm for both H 70/30 and I 70/30 ***(Figure 13C, D)***. The differences in the crystal packing could also impact the kinetics of the dissolution of the Insulin molecule from the crystal; therefore, we performed an *invitro* Insulin release assay of the microcrystalline formulation. From the *invitro* Insulin release study, we found a similar cumulative percentage of Insulin released out from the crystals at 3 hours (~20%) and 13 hours (~33%) ***(Figure 13 E)***.

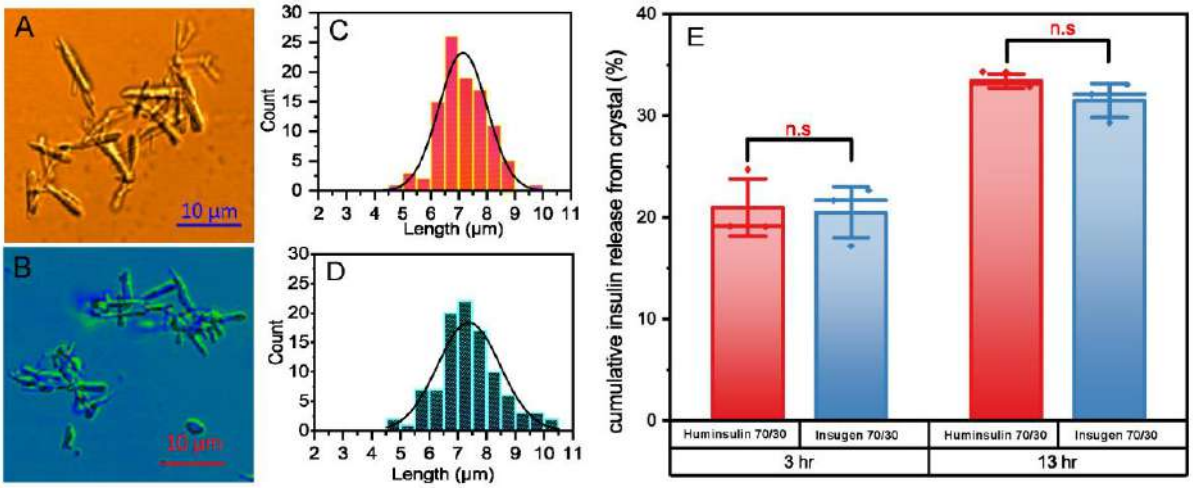

Figure 13. Comparative study of Morphology and crystal release of Humulin and Insugen microcrystal*: Comparison of the Insulin crystal size of H 70/30 (A) and I 70/30 (B) by bright field Microscopy. The length of 100 Insulin crystals from each drug was measured, and a graph was plotted between the length of the crystal (X-axis) and the number of counts (Y-axis). The length of Huminsulin® 70/30 (C) was found in the range of 5-9 µm, and Insugen® 70/30 (D) was in the range of 5.5-10µm. (E) Release of Insulin from the crystal was observed for 3 and 13 hrs in release buffer (phosphate buffer pH 7.4, with 0.01% of tween-80), and a t-test was performed and found no significant difference in both cases.*

The release of Insulin from the crystals under laboratory conditions is comparatively slow. In physiological conditions, enzymes and proteases play a role in the faster release of Insulin from the crystals (Sultan, Mahdi, and Kwon 2020). Nonetheless, despite having differences in the crystal packing, we could not observe any significant difference in the release of protein from the crystal for both drug products.

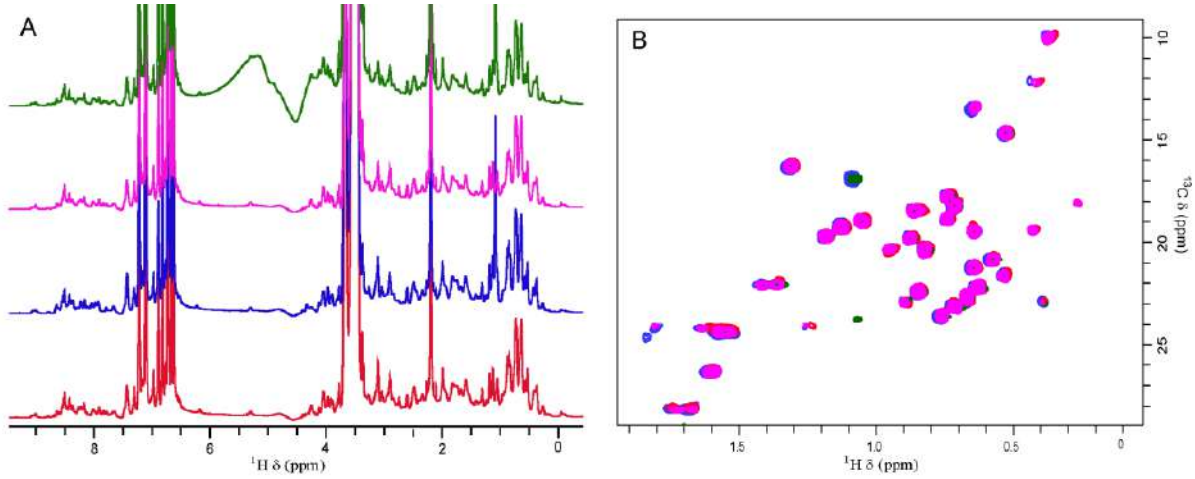

Figure 14. HOS comparison of monomeric Insulin (acidified) of Humulin and Insugen formulation: *Comparison of 1D ^{1}H NMR (A)and 2D ^{1}H-^{13}C SOFAST-HMQC NMR (B)spectra of spectra of Insugen® 70/30 (Red), Huminsulin® 70/30 (Blue), Insugen® 50/50 (magenta) and Huminsulin® 50/50 (green) at pH 2.3. (Monomeric state). 2µl of 6N HCl was added to 500 µl of biphasic suspension drug formulation and 50 µl of deuterium oxide was added for spin lock. All drug suspension was converted to clear solution along with the conversion of higher order hexameric insulin to monomeric form. 2D ^{1}H-^{13}C SOFAST-HMQC NMR spectra were recorded with 1024 × 96 points in the direct and indirect dimensions, respectively with acquisition time of 48 msec. The same data set was zero filled to yield a data matrix of size 4096 × 512 before Fourier transformation.*

To further assess if the differences in the crystal packing have led to perturbation in the active state of Insulin (monomer), we performed a structural assessment of acidified samples of both drug products. For this, the Sample pH was adjusted to 2.3. Next, 1D ^{1}H and 2D ^{1}H-^{13}C SOFAST Methyl HMQC NMR were recorded ***(Figure 14).*** Insulin is known to exist in its monomeric state at lower pH. Both drug products showed a highly comparable spectral profile indicating highly similar monomeric (active) conformation. While variations in the crystal packing, as observed by ssNMR and WAXS studies, didn't show any direct impact on its functionally active state or release, it may have an impact on the long-term storage of the samples. Insulin biphasic suspension products are marketed as multi-dosage vials/pens and can be stored at ambient temperature during the actual usage of typically ~30 days. At ambient temperature, crystal packing differences may lead to differences in the degradation kinetics; hence, it is important to have appropriate monitoring and control during the manufacturing of suspension formulations.

As 1D ^{13}C CPMAS spectra does not have high resolution and also the 2D ^{13}C-^{1}H did not show enough resolution to substantiate any extra atomic level information, we try to develop a selective excitation pulse scheme to obtain the methyl spectra as fingerprint of the protein.

Band selective excitation pulse design

Selective excitation of a particular nuclei or bandwidth has played a crucial role in the development of advanced pulse sequences for spectral manipulation in solution and solid-state NMR. In the realm of solids, selective excitation finds extensive applications in multidimensional spectrum assignment, distance measurements, and overcoming spectral

overlap in uniformly labelled systems. Existing approaches for achieving selective excitation include the use of Gaussian-shaped pulses [74], DANTE (Delay alternating with nutation for tailored excitation) [75,76]), heteronuclear selective phase-optimized recoupling (hetSPR) [77], SELDOM (SELectivity by Destruction of Magnetization) [78], or computer-optimized E-family sequences [79]. In this study, we introduce a reflection symmetric version of the DANTE sequence with variable flip angles followed by a hard π pulse and a specific delay time for band selective excitation ***(Figure 15).*** This allowed the selection of desirable aliphatic side-chain resonances in the ^{13}C-^{1}H HETCOR spectra with enhanced resolution.

A comprehensive theoretical formulation of this sequence

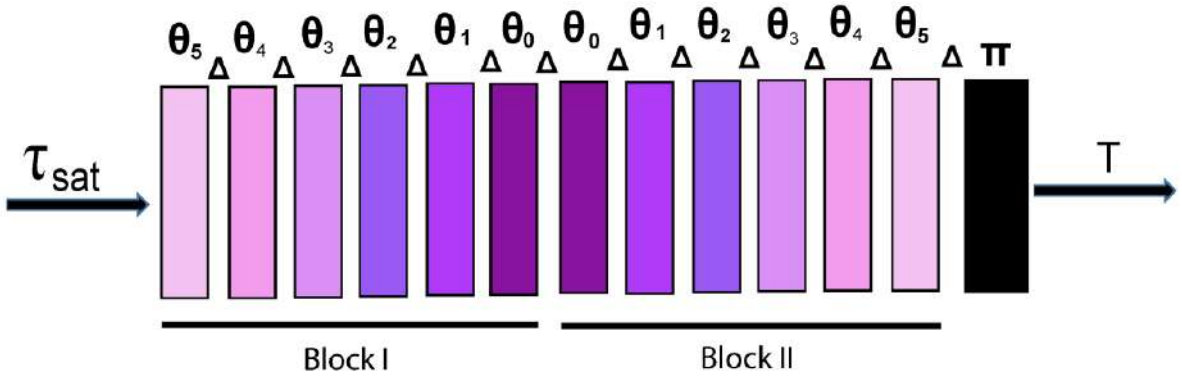

Figure 15. Scheme of Reflection Symmetric DANTE*. (where, Δ (delay)= 26.31 μs, Θ_0= π* 0.0208, Θ_1= π* 0.0412, Θ_2= π* 0.0398, Θ_3= π* 0.0375, Θ_4= π* 0.0345, Θ_5= π* 0.0307, π=180° hard pulse in carbon channel)*

Let Δ be rotor period. We design a reflection symmetric Dante sequence with flip angles θ_0, θ_1…, θ_N with the sequences written as $\theta_N - \Delta - \cdots - \theta_1 - \Delta - \theta_0 - \Delta - \theta_0 - \Delta - \cdots - \theta_N - \Delta$ followed by a hard π pulse and delay of time T = (N +1) Δ. N is chosen such that if bandwidth of excitation is B. Then BΔ = 2 π/N. To determine θ_k, we consider the problem of band selective excitation. Consider the evolution of the Bloch vector X (We use Ω_α to denote the rotation matrix, such that α ϵ {x, y, z}) of a spin 1/2, in a rotating frame, rotating around z-axis at Larmor frequency.

$$\frac{dX}{dt} = (\omega\Omega z + A(t)cos\,\theta(t)\Omega x + A(t)sin\,\theta(t)\Omega y)X, \qquad (1)$$

Where A(t) and θ(t) are amplitude and phase of rf-pulse, and we normalize the chemical shift, ω ϵ [-1,1]. In what follows, we choose sin θ(t)=0 and let

$$\frac{dX}{dt} = (\omega\Omega z + u(t)\,\Omega x)X, \qquad (2)$$

where u(t) is the amplitude modulated pulse for t ϵ [0, T].

Going into the interaction frame of chemical shift, using

$$Y(t) = \exp(-\omega\left(t - \frac{T}{2}\right)\Omega z)X(t)$$

We obtain,

$$\frac{dY}{dt} = u(t)(cos(\omega\left(t-\frac{T}{2}\right)\Omega x - sin\left(\omega\left(t-\frac{T}{2}\right)\Omega y\right)Y,$$

$$Y(0) = \exp(\omega\frac{T}{2}\Omega z)X(0) \quad (3)$$

We design u(t) such that for all ω ϵ [-1,1], we have

$$\int_0^T u(t)\cos\left(\omega\left(t-\tfrac{T}{2}\right)\right)dt \sim \theta; \quad \int_0^T u(t)\sin\left(\omega\left(t-\tfrac{T}{2}\right)\right)dt = 0 \qquad (4)$$

Divide [0, T] in intervals of step Δ, over which u(t) is constant. Call these amplitudes, u_{-M} …, u_{-k} …u_0 over [0, T/2], and u_0 …, u_k …u_M over [T/2, T].

$$\int_0^T u(t)\cos\left(\omega\left(t-\frac{T}{2}\right)\right)dt \sim (u_0 + \sum_{k=-M}^{M} u_k\cos(\omega k\Delta)\Delta) \qquad (5)$$

where write Δ = π/N and choose $u_k = u_{-k}$. This insures that sine equation in Eq. (4) above is automatically satisfied. Then we get,

$$\int_0^T u(t)\cos\left(\omega\left(t-\frac{T}{2}\right)\right)dt \sim 2\sum_{k=0}^{M} u_k\cos(\omega k\Delta)\Delta = 2\sum_{k=0}^{M} u_k\cos(kx)\Delta \qquad (6)$$

where for x ϵ [−π/N, π/N] , we have $2\sum_{k=0}^{M} u_k\cos(kx)\Delta \sim \theta$ and 0 for x outside this range. This is a Fourier series, and we get the Fourier coefficients as,

$$u0 = \frac{\theta}{2\pi}; \qquad uk = \frac{2\theta}{\pi}\frac{sin\frac{k\pi}{N}}{\frac{2k\pi}{N}} \qquad (7)$$

For$\theta = \frac{\pi}{2}$, we get,

$$u0 = \frac{1}{4}; \qquad uk = \frac{sin\frac{k\pi}{N}}{\frac{2k\pi}{N}} \qquad (8)$$

Let $\theta_k = u_k\Delta$ $\theta k = uk$, then the sequence is implemented as a Dante sequence. We choose M = N. Starting from the initial state

$$X(0) = \begin{matrix}0\\0\\1\end{matrix}$$

, we have from equation (3),

$$X(T) \sim \exp\left(\omega\frac{T}{2}\Omega_z\right)\exp\left(\int_0^T u(t)\cos\left(\omega\left(t-\frac{T}{2}\right)\right)dt\Omega_x\right)\exp\left(\omega\frac{T}{2}\Omega_z\right)X(0) \sim \exp\left(\omega\frac{T}{2}\Omega_z\right)\exp(\pi\Omega_x)X(0)$$

for $\omega\Delta \epsilon [-\pi/N, \pi/N]$. There is no excitation outside the desired band.

This state is dephased on the Bloch sphere equator. To refocus this, we need a π pulse and delay in the end of the sequence.

Testing the performance of selective excitation on microcrystalline SH3 protein

To showcase the effectiveness of the selective excitation scheme, we conducted an initial test on a microcrystalline protein called alpha-SH3 domain, which was uniformly labeled with ^{13}C and ^{15}N isotopes. The two reflection symmetric excitation blocks of DANTE were integrated into the ^{13}C channel of a 2D HETCOR experiment, with proton detection, as outlined in the pulse program ***(Figure 16A).*** By employing the reflection symmetric DANTE sequence on the ^{13}C nuclei, we were able to selectively excite the narrow ^{13}C region encompassing methyl, methine, and methylene side chain resonances (0 to ~30 ppm) while effectively suppressing unwanted signals from other regions, as depicted in ***Figure 16B***. For comparison, the one-dimensional ^{13}C spectra of the aliphatic region were obtained using a 90^0 hard pulse broadband excitation. Additionally, we conducted a 2D ^{1}H detected ^{13}C-^{1}H HETCOR version of the selective excitation experiment and the corresponding results are presented in ***Figure 16C.***

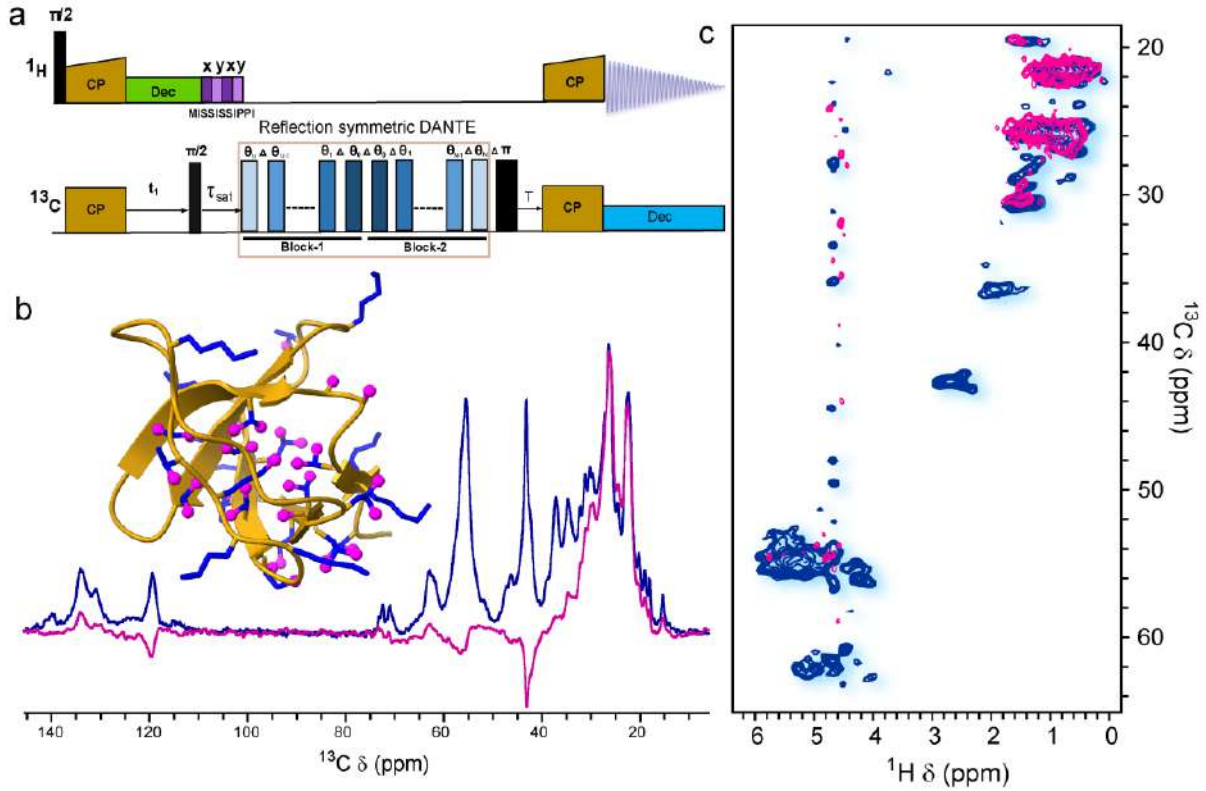

Figure 16. Testing the performance of selective excitation on microcrystalline SH3 protein*: (A) Pulse diagram to acquire methyl and aliphatic side-chain fingerprint employing 1H detection. (B) One-dimensional 13C spectra of microcrystalline SH3 domain showing the performance of selective excitation scheme (pink) and broadband excitation (blue), (C) Overlay spectra of 2D 13C-1H CP-HETCOR without reflection symmetric DANTE block incorporated with 1H detection (blue) and by using the novel pulse scheme showed in Figure 2 A (pink).*

The spectral profile of the newly designed pulse scheme shows a near-perfect overlap with the traditional 2D-^{13}C-^{1}H HETCOR experiment in the methyl and Methylene region (0 to ~30 ppm). However, no significant enhancement in resolution could be obtained. At a high field (750 MHz) and high abundance of isotopically labelled ^{13}C, strong dipolar coupling induced anisotropic effects are likely to have contributed to line broadening and poor resolution. Also, the inherent heterogeneous nature of the sample cannot be ruled out. This inclined us to test our hypothesis on suspension formulations of human insulin, which is present in its natural abundant state and is well studied to exist in a microcrystalline state, with crystal size varying between 3-8 microns [24] These crystals are well studied by X-ray crystallography and reported to have a single hexameric unit per crystal and remaining solvent [23].

3.4.3 Resolution enhancement in insulin suspensions at natural abundance

In the previous study conducted by Zhou et al., the viability of proton-detected 2D ^{1}H-^{13}C and ^{1}H-^{15}N CP-HETCOR techniques on biopharmaceuticals at their natural isotopic abundance was demonstrated [41]. This study focused on the backbone (-CH) region of the ^{13}C-^{1}H HETCOR spectrum, providing valuable information regarding the overall folding pattern of protein-based pharmaceuticals. In the current study, our focus was primarily on the sensitivity and resolution enhancement in the methyl and aliphatic side-chain regions of the spectrum, which offer insights into the secondary, tertiary, and quaternary structural features collectively referred to as the higher-order structure (HOS).

Proton-detected 2D ^{13}C-^{1}H HETCOR experiment at Carbon-13 natural abundance can sometimes take a significant amount of time to achieve a decent signal-to-noise (S/N) ratio as compared to uniformly labeled protein samples. In order to significantly reduce on acquisition time, we incorporated non-uniform sampling (NUS), which is quite common in recording multi-dimensional spectra in biomolecular NMR [80–82]. However, it is not widely used in pharmaceutical solid-state NMR applications where the low natural abundance of NMR active isotopes is usually a limiting factor. Over the years, several NUS schemes have been proposed to sample time domain signals more frequently when the signal is strong, and less when it is weak [83,84]. There are only a few studies employing solid-state NMR where a detailed analysis of NUS schemes has been described [85–87]. Conventional data processing methods such as Fast Fourier Transformation (FFT) and Linear Prediction (LP) cannot be applied to data collected by NUS. Instead, different algorithms, including Maximum Entropy, Maximum Likelihood Method (MLM), multi-dimensional decomposition (MDD), and Compressed Sensing (CS), are used for the reconstruction of NUS data. For the current study, we opted for Sinusoidal Poisson-gap sampling scheme for acquisition and compressed sensing for reconstruction and processing that has been shown to enhance the resolution of the spectra with high fidelity [88–90].

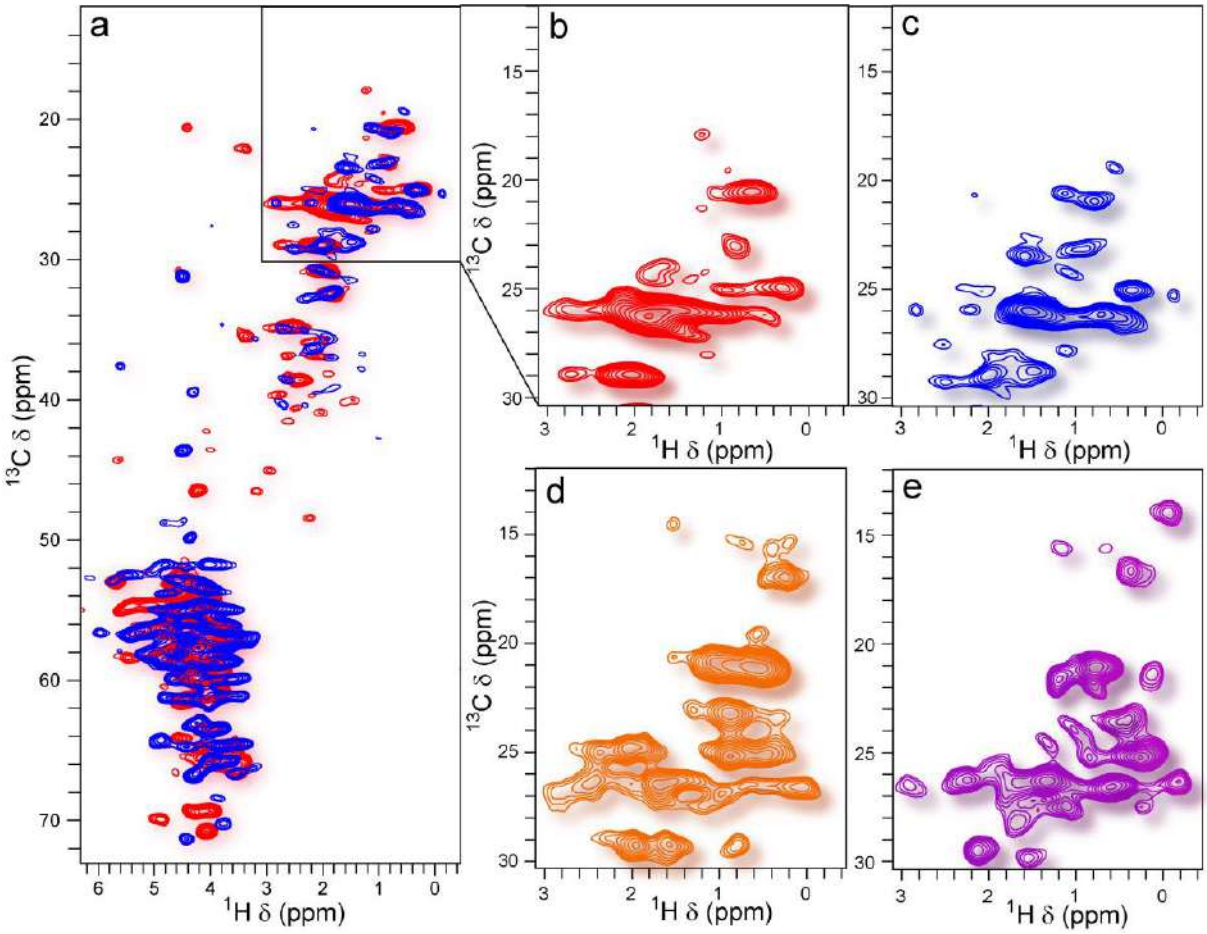

Figure 17. Resolution enhancement in insulin suspensions at natural abundance: *(A) 2D ^{13}C-^{1}H CP-HETCOR overlay of Humulin N (red) and Humulin 70/30 (blue) with ^{1}H detection (B, C) Zoomed region from the overlay spectra of figure 3 A. (D, E) 2D ^{13}C-^{1}H CP-HETCOR with ^{1}H detection using the pulse sequence showed in Figure1 A. (with the reflection symmetric DANTE block) of Humulin N (orange) and Humulin 70/30 (magenta) respectively.*

We performed ^{1}H-detected 2D ^{13}C-^{1}H HETCOR experiments on representative batches of Humulin® N (100% suspension) and Humulin® 70/30 (biphasic mix of 70 parts of suspension human insulin and 30 parts of soluble human insulin) and observed a significant spectral overlap between both ***(Figure 17A),*** indicating comparable HOS of human insulin in the two drug products despite minor differences in the formulations. Subsequently, we repeated the experiment with the reflection symmetry DANTE incorporated in the ^{1}H-detected 2D ^{13}C-^{1}H HETCOR on both the samples ***(Figure 17D, E),*** where we have selectively excited the methyl and aliphatic regions (0-30 ppm). By adjusting the magnetization transfer scheme in the Heteronuclear correlation (HETCOR) experiment, and reconstruction of NUS data by a compressed sensing algorithm, it is possible to obtain information from either the mobile or rigid components [91]. The spectral profile, thus obtained, in the selected region, showed improvement in resolution and increase in S/N about ~2-2.5-fold ***(Figure 18),*** which is noticeable from the appearance of flexible methyl peaks (below 20 ppm).

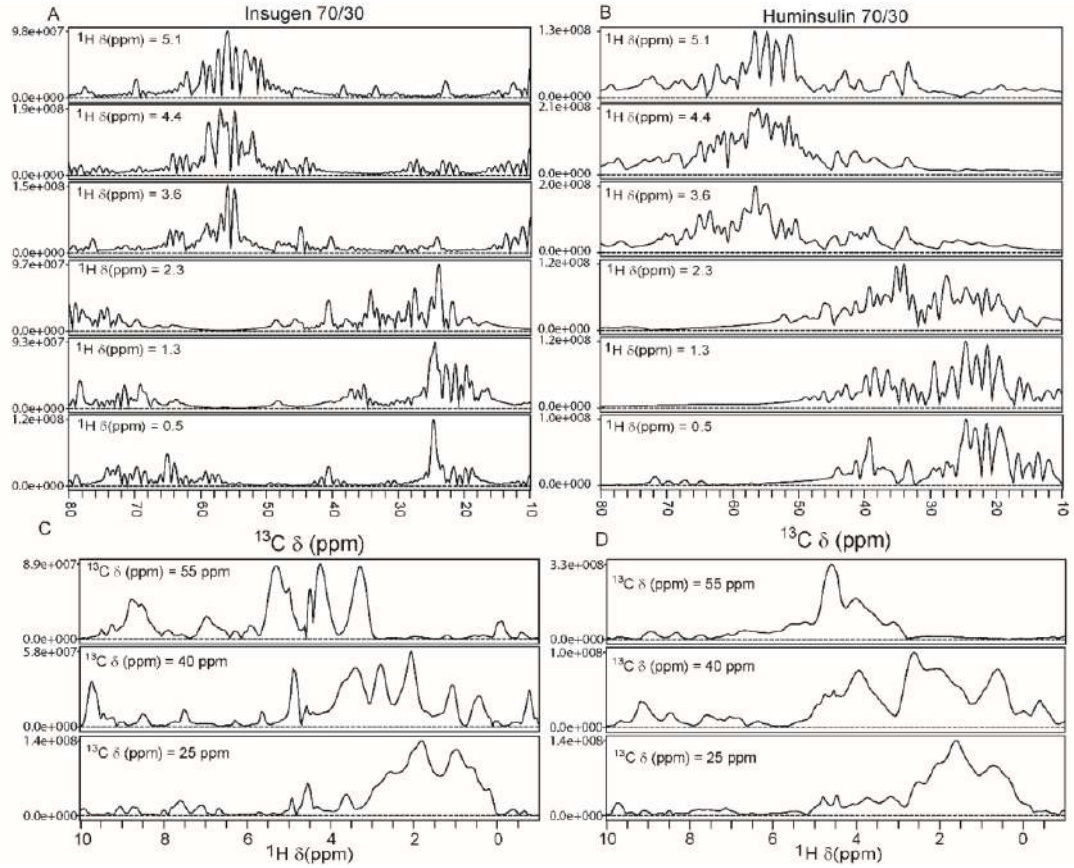

Figure 18: Sensitivity enhancement in selective excitation scheme: *Signal-to-noise (sensitivity) comparison of (A) Humulin N non-selective excitation (B) Humulin 70/30 non-selective excitation (C) Humulin N selective excitation (D) Humulin 70/30 selective excitation*

The observed line width of Humulin® N decreased from 305 Hz to 188 Hz (~40% reduction) and for Humulin® 70/30 it decreased from 182 Hz to 149 Hz (~20% reduction) ***(Figure 19).*** The enhanced resolution and sensitivity allowed us to investigate the higher-order structure of insulin in the microcrystalline suspension without perturbing the crystalline state of samples and preserving the overall higher-order structure.

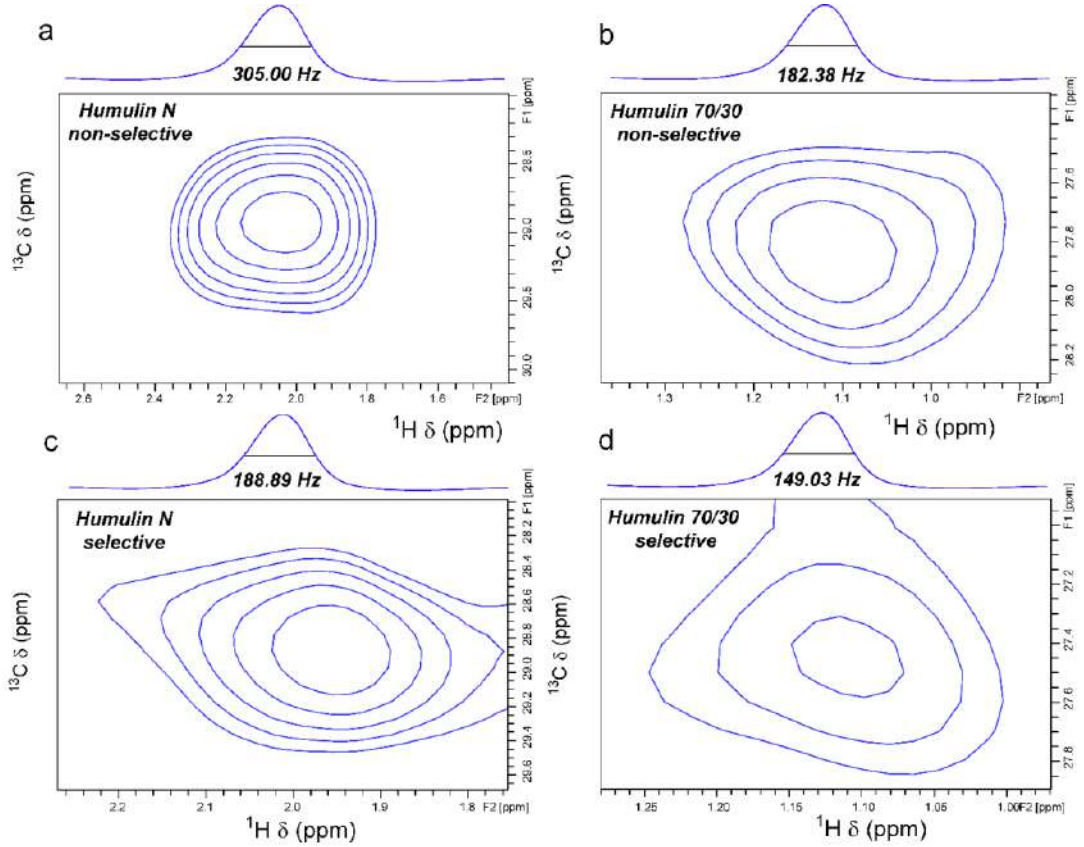

Figure 19. Line width comparison in selective vs non-selective excitation scheme: *Line width comparison of (a) Humulin N non-selective excitation (b) Humulin 70/30 non-selective excitation (c) Humulin N selective excitation (d) Humulin 70/30 selective excitation*

3.4.4 Allosteric transition of hexameric insulin

The intrinsic property of insulin is to self-associate, which plays a key role in therapeutic formulations. In particular, the hexameric state is promoted by the presence of phenolic ligands- and Zn^{2+} (or other divalent transition metals) that exhibit greater stability than insulin dimers and monomers and hence are needed for a longer shelf-life [62]. Metal-bound insulin hexamers also exhibit phenol-induced allostery [92–94]. Insulin can co-exist in equilibrium as a mixture of all taut (T_6), all relaxed (R_6), or half taut-half relaxed (T_3R_3) protein conformations [95–98]***(Figure 21a).*** The structural difference between T-state and R-state is only in the first 8 amino-acids of B-chain of insulin monomer, where B1-B8 folds into helical state in R-state in the presence of phenolic ligands, while they remain in extended conformation in T-state. The T-R transition has been reported by using different biophysical techniques such as Circular dichroism (CD) [99], FTIR [100], and Raman spectroscopy[101], and the representative crystal states are available in the literature [102–104]. Although the physiological significance of the T- and R-states is yet unknown, it has been proposed that the R-state is primarily associated with the higher stability of insulin in its formulated state. The T-state is crucial for correct proinsulin and/or insulin folding [94,105] and is biologically more relevant. However, due to limited structural information, insulin Receptor binding does not always entail the T or the R-state [106,107]. The two hexameric states, T_6 and R_6, differ significantly in their dissociation kinetics, with the T_6 state prevailing in the physiological settings and undergoing dissociation to dimers within a

few minutes, while the R_6 state has a significantly longer lifetime ranging from hours to days [108,109]. To increase the shelf life of therapeutic formulations, phenol is often added as an excipient as it tilts the equilibrium towards the more long-lasting R_6 state [110]. This is expected to decrease dissociation and, subsequently, prevent the occurrence of off-pathway equilibria. Existence of multiple hexameric states observed in solution as well as in amorphous and crystalline states is a clear link to occurrence of phenol-bound allostery [100,111,112].

With improved sensitivity and resolution of methyl peaks, we were able to discern mechanistic snapshots from the representative spectra of Humulin® N and Humulin® 70/30. Structural features were deduced from the chemical shift predictions for the insulin hexamer with the help of an online server called SHIFTX2 [57] (Table A5). The predicted and the experimental spectra showed very good correlation, that allowed putative assignment of a few important peaks, representative of different conformational states of human insulin. An up-field shifted peak in the ω_1-dimension (^{1}H: -0.175 ppm) was observed in selective methyl spectra of Humulin® 70/30 (***Figure 21e***; labelled in cyan). The phenolic ligand in the R_6 hexamer largely resides in a hydrophobic pocket at the dimer-dimer interface, which is formed by HisB5, LeuB11, and residue CysA6, CysA11 [110] ***(Figure 21b)*** from different dimer subunits. As methyl of residue LeuB11 is surrounded by high hydrophobic groups, the corresponding resonance is expected to show an up-field shift. The peak below zero ppm was therefore assigned as LeuB11 and is also labelled as a marker for the Insulin in predominantly R_6 hexamer conformation, as in a T_6 or T_3R_3 state, the dimer interface would be less crowded, and therefore an up-field shift is not expected. Interestingly, this peak was found missing and/or shifted downfield in the Humulin® N spectrum ***(Figure 21f)*** indicative of an alternative hexameric state(s) (T_6 or T_3R_3). In the T-state, enhanced flexibility of B1-B8 residues ***(Figure 21d)*** would lead to a decrease in the magnetization transfer via dipolar couplings, which will be reflected as a decrease in the intensity of the corresponding peaks in the B1-B8 residues. This was well corroborated as the decreased intensity of the LeuB6 peak in the Humulin® N spectrum, indicating LeuB6 to be another marker peak for the identification of enhanced T-state contributions in the hexamer. Further, A weighted average combination of different ratios of PDB structures in the absence (T6 conformation; PDB ID: 1MSO) and presence (R6 conformation; PDB ID: 1ZNJ) of phenolic ligand was used for back prediction of chemical shifts. A comparatively better match with 95% of R6:1ZNJ and 5% of T6:1MSO was observed for Humulin® 70/30 and 75% of R6:1ZNJ and 25% of T6:1MSO was observed for Humulin® N ***(Figure 20).***

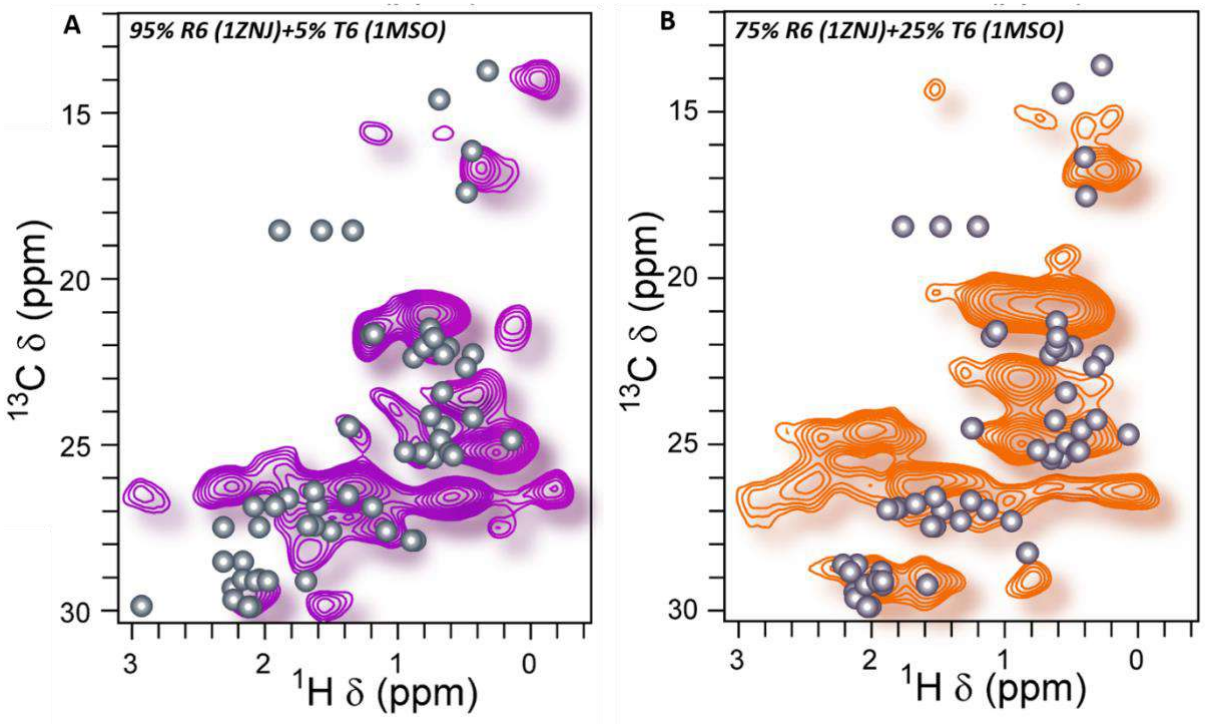

Figure 20. Overlay of experimental vs predicted NMR spectra of Insulin: *Overlay of 2D ^{13}C-^{1}H HETCOR with ^{1}H detection using the reflection symmetric DANTE pulse sequence of (A) Humulin 70/30 (magenta) and weighted average of predicted spectra of PDB: 95% 1ZNJ (R6) and 5% 1MSO (T6) from SHIFTX2 (B) Humulin N (orange) and weighted average of predicted spectra of PDB: 75% 1ZNJ (R6) and 25% 1MSO (T6) from SHIFTX2*

To the best of our knowledge, this is the first NMR study, which could demonstrate evidence of different hexameric states in the insulin suspensions using methyl peak markers. The conformational fluctuations amongst different hexameric states is a well-known phenomenon as is also demonstrated by single crystal X-ray studies, where each of the three states could be crystallized [102–104]. These conformational fluctuations happen due to the dynamic movement ('in' and 'out' of the hexamer) of phenolic ligands present in the drug formulations. These 'in' and 'out' movements have been studied as 'ligand escaping pathways' using computational biology tools, primarily using molecular dynamic simulations[108,113].

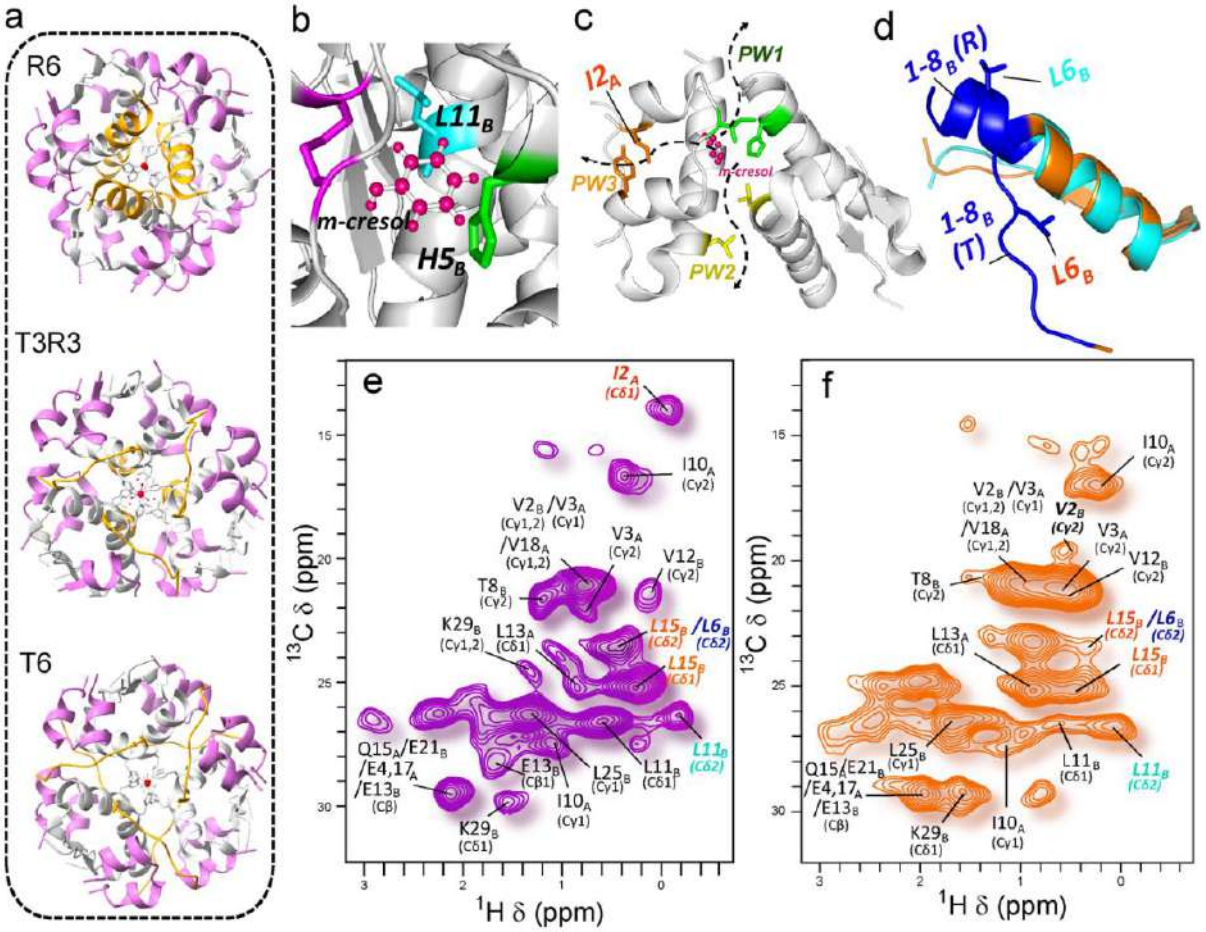

Figure 21. Different conformational transition of hexameric insulin: *(a) Different hexameric conformation of insulin; Insulin R6 (PDB: 1ZNJ), Insulin T3R3 (1TRZ), Insulin T6 (1MSO). (b) Hydrophobic pocket of R form of insulin involving CysA6, CysA11 (both marked in purple), HisB5 (green) and Leu11 of B-chain (cyan). (c) Different phenol escape pathways named PW1, PW2, PW3 and the respective gatekeeper residues are marked IleA10, HisB5 (PW1: green), LeuA13, LeuB17 (PW2: yellow), IleA2, TyrA19(PW3: orange). (d) superimpose of B-chain of insulin R form (cyan) and T form (orange), first 1-8 amino acid of both the form marked in blue. (e) Most probable resonance assignment of 2D ^{13}C-^{1}H HETCOR with the reflection symmetric DANTE block of Humulin 70/30 (chemical shift predicted from SHIFTX2), (f) Most probable resonance assignment of 2D ^{13}C-^{1}H HETCOR with the reflection symmetric DANTE block of Humulin® N (chemical shift predicted from SHIFTX2).*

The studies by Vashisht et al. [113] and Adam et al. [108] revealed the existence of three distinct escaping pathways (named PW1, PW2, and PW3) of phenol from its hydrophobic core ***(Figure 21c).*** In PW1, the process of phenol escape is accomplished via a gate-pushing mechanism, wherein the phenol molecule forces its way through the "gatekeeper" residues IleA10 and HisF5. On the other hand, PW2 involves a gate-hopping mechanism whereby the phenol molecule utilizes an existing escape channel located between IleA10/HisF5 and LeuA13/LeuH17 (Insulin hexameric sub-units are named from A-F). PW3 follows a different pathway, as the phenol molecule escapes through a densely packed region in the insulin monomer's core, primarily comprised of chains A and B, and is predominantly lined by specific

residues such as IleA2, GlnA5, CysA6, CysA11, LeuA16, TyrA19, LeuB11, and LeuB15. The two residues IleA2 and TyrA19 act as gatekeeper residues for PW3 [108]. Through MD simulations, Adam et al. have also shown an increase in solvation of the phenol and gatekeeper residues during phenol escape from its hydrophobic core [108]. This increase in the hydration level of the IleA2 (gatekeeper residue of PW3) will potentially lead to an increase in the flexibility of IleA2 sidechains. Interestingly, the intensity of the IleA2 peak is minimal, as observed in Humulin® N indicative of the release of phenolic ligand via the PW3 pathway. This is further corroborated by observing the change in the intensity of the LeuB15 in Humulin® N in comparison to Humulin® 70/30 spectra. In total, the enhanced resolution in the indirect dimension has provided sufficient information to deduce multiple transitions involved in the insulin hexameric states. It is to be noted that slightly different conformations observed in the two insulin suspension formulations represent the dynamic picture of the insulin hexameric states due to re-arrangements of phenolic ligands. While these differences may have some impact on product stability, these differences are not expected to have any meaningful clinical differences as the biologically active moiety in insulin is its monomer. Nonetheless, our novel experimental scheme allowed discerning different states of insulins, and we believe it can be extended to other biopharmaceuticals, where different states may lead to safety or immunogenicity concerns. The proton-detection MAS with NUS sampling scheme proposed in this work offers a high-resolution structural approach to comprehensively differentiate various conformations between two formulations of biopharmaceuticals. The signature obtained in a residue-type or at chemical moiety level exceed the capabilities of other analytical methods such as powder X-ray diffraction or SAXS, typically used for microcrystalline or amorphous samples. The diffraction patterns obtained from these techniques for large microcrystalline proteins/mAb therapeutics generate limited structural information. Therefore, we believe that the proposed method can serve as an advanced tool in drug development, investigating protein in-stability events, and quality control.

Discussion

In this work, we have proposed a combinatorial approach for higher-order structure similarity comparison for biphasic microcrystalline suspension in its formulated state. We separated the soluble and solid fractions from the biphasic microcrystalline formulation and performed 1D ^{1}H liquid and 1D ^{13}C CP-MAS NMR spectra for both solution (30% part) and the suspension (70% part) part, respectively. For qualitative HOS assessment, we compare the 1D NMR spectra independently for both solution and a solid fraction, followed by PCA and Mahalanobis distance for quantitative assessment. 1D ^{13}C CP-MAS NMR spectra were sufficient to provide information regarding the rigidity of the molecule (confirmed via hC-CP build-up curves), and crystal packing (corroborates data from WAXS) which is reflected by the intensity differences in the CP transfer. Therefore, we strongly suggest, a combination of 1D ^{1} H liquid NMR and 1D ^{13} C CP MAS NMR experiments is an elegant tool to compare the HOS and dynamics of the biphasic/microcrystalline suspension formulation. However 2D HETCOR experiments did not provide any critical information due to low concentration and low resolution for the solution and solid fractions, respectively. Next we tried to develop a technique for obtaining methyl and aliphatic side-chain spectral fingerprints with the site-specific resolution of

pharmaceutical insulin suspensions. The combination of a novel selective excitation scheme, proton detection at fast MAS, and non-uniform sampling provides mechanistic insights into the higher-order structure of pharmaceutical insulin. The use of non-uniform sampling improved the signal-to-noise ratio and sensitivity of multiple peaks, achieving a resolution in the indirect dimension that was not achievable with traditional uniform acquisition at natural isotopic abundance. The enhanced resolution and sensitivity allowed us to experimentally observe different hexameric states in insulin formulations. Additionally, the current study identified the ligand (Phenol) escape route from the protein complex (insulin hexamer) on the basis of change in intensities observed for key marker residues. To the best of our knowledge, this is the first detailed experimental evidence of the key residues involved in a phenol escaping pathway leading to conformational transitions to different hexameric states in insulins. This approach can be seamlessly extended to crystalline suspensions of monoclonal antibodies (mAbs), which are being considered as a potential alternative formulation strategy for subcutaneous delivery at high concentrations[64]. As the molecular weight increases, the resolution tends to decrease; nevertheless, our approach allows for a notably improved enhancement in both sensitivity and resolution and therefore structural details of complex biopharmaceuticals can be deduced in a precise manner. In summary, the method we presented can be applied to a wide range of protein pharmaceuticals in suspension state, allowing for the determination of atomic-level higher-order structural information previously not attainable in their native formulation state.

Impact of the research in the advancement of knowledge or benefit to mankind

- The initial technique we developed serves as a sophisticated tool for comparing the higher-order structure (HOS) and dynamics of biphasic/microcrystalline suspension formulations.

- Our second method, characterized by high resolution, is applicable across a broad spectrum of protein pharmaceuticals in suspension. It enables the determination of atomic-level structural details that were previously unattainable in their native formulation state.

- Previously employed methods for determining the HOS of microcrystalline formulations were either low resolution or necessitated perturbation of the native formulation, potentially resulting in structural alterations to biopharmaceuticals.

- Furthermore, our approach facilitates the investigation of batch-to-batch variation during manufacturing and storage, as well as enables biosimilarity comparisons in biphasic/microcrystalline suspensions, thereby enhancing product quality.

- Our approach can assist regulatory bodies such as CDSCO, USFDA, and EMA by offering a sophisticated tool to monitor the quality of crystalline biopharmaceuticals in their native state, a capability previously unachievable. This represents a significant advancement in monitoring critical quality attributes of biopharmaceuticals.

References

(1) Mathaes, R.; Koulov, A.; Joerg, S.; Mahler, H.-C. Subcutaneous Injection Volume of Biopharmaceuticals-Pushing the Boundaries. *J Pharm Sci* **2016**, *105* (8), 2255–2259. https://doi.org/10.1016/j.xphs.2016.05.029.

(2) Jackisch, C.; Müller, V.; Maintz, C.; Hell, S.; Ataseven, B. Subcutaneous Administration of Monoclonal Antibodies in Oncology. *Geburtshilfe Frauenheilkd* **2014**, *74* (4), 343–349. https://doi.org/10.1055/s-0034-1368173.

(3) Joubert, M. K.; Luo, Q.; Nashed-Samuel, Y.; Wypych, J.; Narhi, L. O. Classification and Characterization of Therapeutic Antibody Aggregates. *J Biol Chem* **2011**, *286* (28), 25118–25133. https://doi.org/10.1074/jbc.M110.160457.

(4) Liu, J.; Nguyen, M. D. H.; Andya, J. D.; Shire, S. J. Reversible Self-Association Increases the Viscosity of a Concentrated Monoclonal Antibody in Aqueous Solution. *J Pharm Sci* **2005**, *94* (9), 1928–1940. https://doi.org/10.1002/jps.20347.

(5) Lilyestrom, W. G.; Yadav, S.; Shire, S. J.; Scherer, T. M. Monoclonal Antibody Self-Association, Cluster Formation, and Rheology at High Concentrations. *J Phys Chem B* **2013**, *117* (21), 6373–6384. https://doi.org/10.1021/jp4008152.

(6) Luo, H.; Lee, N.; Wang, X.; Li, Y.; Schmelzer, A.; Hunter, A. K.; Pabst, T.; Wang, W. K. Liquid-Liquid Phase Separation Causes High Turbidity and Pressure during Low PH Elution Process in Protein A Chromatography. *J Chromatogr A* **2017**, *1488*, 57–67. https://doi.org/10.1016/j.chroma.2017.01.067.

(7) Nishi, H.; Miyajima, M.; Nakagami, H.; Noda, M.; Uchiyama, S.; Fukui, K. Phase Separation of an IgG1 Antibody Solution under a Low Ionic Strength Condition. *Pharm Res* **2010**, *27* (7), 1348–1360. https://doi.org/10.1007/s11095-010-0125-7.

(8) Reiche, K.; Hartl, J.; Blume, A.; Garidel, P. Liquid-Liquid Phase Separation of a Monoclonal Antibody at Low Ionic Strength: Influence of Anion Charge and Concentration. *Biophys Chem* **2017**, *220*, 7–19. https://doi.org/10.1016/j.bpc.2016.08.003.

(9) Cohen, S. L.; Price, C.; Vlasak, J. Beta-Elimination and Peptide Bond Hydrolysis: Two Distinct Mechanisms of Human IgG1 Hinge Fragmentation upon Storage. *J Am Chem Soc* **2007**, *129* (22), 6976–6977. https://doi.org/10.1021/ja0705994.

(10) Moritz, B.; Stracke, J. O. Assessment of Disulfide and Hinge Modifications in Monoclonal Antibodies. *Electrophoresis* **2017**, *38* (6), 769–785. https://doi.org/10.1002/elps.201600425.

(11) Basu, S. K.; Govardhan, C. P.; Jung, C. W.; Margolin, A. L. Protein Crystals for the Delivery of Biopharmaceuticals. *Expert Opin Biol Ther* **2004**, *4* (3), 301–317. https://doi.org/10.1517/14712598.4.3.301.

(12) Merkle, H. P.; Jen, A. A Crystal Clear Solution for Insulin Delivery. *Nature biotechnology*. United States August 2002, pp 789–790. https://doi.org/10.1038/nbt0802-789.

(13) Shenoy, B.; Wang, Y.; Shan, W.; Margolin, A. L. Stability of Crystalline Proteins. *Biotechnol Bioeng* **2001**, *73* (5), 358–369. https://doi.org/10.1002/bit.1069.

(14) Yang, M. X.; Shenoy, B.; Disttler, M.; Patel, R.; McGrath, M.; Pechenov, S.; Margolin, A. L. Crystalline Monoclonal Antibodies for Subcutaneous Delivery. *Proc Natl Acad Sci U S A* **2003**, *100* (12), 6934–6939. https://doi.org/10.1073/pnas.1131899100.

(15) Judge, R. A.; Forsythe, E. L.; Pusey, M. L. The Effect of Protein Impurities on Lysozyme Crystal Growth. *Biotechnol Bioeng* **1998**, *59* (6), 776–785. https://doi.org/10.1002/(sici)1097-0290(19980920)59:6<776::aid-bit14>3.0.co;2-3.

(16) Dos Santos, R.; Carvalho, A. L.; Roque, A. C. A. Renaissance of Protein Crystallization and Precipitation in Biopharmaceuticals Purification. *Biotechnol Adv* **2017**, *35* (1), 41–50. https://doi.org/10.1016/j.biotechadv.2016.11.005.

(17) Brange, J.; Vølund, A. Insulin Analogs with Improved Pharmacokinetic Profiles. *Adv Drug Deliv Rev* **1999**, *35* (2–3), 307–335.

(18) Markussen, J.; Hougaard, P.; Ribel, U.; Sørensen, A. R.; Sørensen, E. Soluble, Prolonged-Acting Insulin Derivatives. I. Degree of Protraction and Crystallizability of Insulins Substituted in the Termini of the B-Chain. *Protein Engineering, Design and Selection* **1987**, *1* (3), 205–213.

(19) Sultan, M. H.; Mahdi, W. A.; Kwon, Y. M. Insulin Release from NPH Insulin-Loaded Pluronic® F127 Hydrogel in the Presence of Simulated Tissue Enzyme Activity. *Processes* **2020**, *8* (10), 1320.

(20) Clogston, L. C.; Christian, R. T.; Osslund, D. T.; Freeman, E. Sclerostin Antibody Crystals and Formulations Therof, 2016.

(21) Fraunhofer, W.; Borhani, W. D.; Winter, G.; Gottschalk, S. Compositions and Methods for Crystallizing Antibodies.

(22) Quinternet, M.; Starck, J.-P.; Delsuc, M.-A.; Kieffer, B. Heteronuclear NMR Provides an Accurate Assessment of Therapeutic Insulin's Quality. *J Pharm Biomed Anal* **2013**, *78–79*, 252–254. https://doi.org/10.1016/j.jpba.2013.02.016.

(23) Norrman, M.; Ståhl, K.; Schluckebier, G.; Al-Karadaghi, S. Characterization of Insulin Microcrystals Using Powder Diffraction and Multivariate Data Analysis. *J Appl Crystallogr* **2006**, *39* (3), 391–400.

(24) Pujahari, S. R.; Mali, P. S.; Purusottam, R. N.; Kumar, A. Combined Liquid-State and Solid-State Nuclear Magnetic Resonance at Natural Abundance for Comparative Higher Order Structure Assessment in the Formulated-State of Biphasic Biopharmaceutics. *Anal Chem* **2023**. https://doi.org/10.1021/acs.analchem.2c05485.

(25) Søeborg, T.; Rasmussen, C. H.; Mosekilde, E.; Colding-Jørgensen, M. Absorption Kinetics of Insulin after Subcutaneous Administration. *European journal of pharmaceutical sciences* **2009**, *36* (1), 78–90.

(26) Søeborg, T.; Rasmussen, C. H.; Mosekilde, E.; Colding-Jørgensen, M. Bioavailability and Variability of Biphasic Insulin Mixtures. *European journal of pharmaceutical sciences* **2012**, *46* (4), 198–208.

(27) Pechenov, S.; Shenoy, B.; Yang, M. X.; Basu, S. K.; Margolin, A. L. Injectable Controlled Release Formulations Incorporating Protein Crystals. *J Control Release* **2004**, *96* (1), 149–158. https://doi.org/10.1016/j.jconrel.2004.01.019.

(28) Schellekens, H. Biosimilar Therapeutics-What Do We Need to Consider? *NDT Plus* **2009**, *2* (Suppl_1), i27–i36. https://doi.org/10.1093/ndtplus/sfn177.

(29) Mohammad, M. A.; Grimsey, I. M.; Forbes, R. T. Mapping the Solid-State Properties of Crystalline Lysozyme during Pharmaceutical Unit-Operations. *J Pharm Biomed Anal* **2015**, *114*, 176–183.

(30) Sage, J. T.; Zhang, Y.; McGeehan, J.; Ravelli, R. B. G.; Weik, M.; Van Thor, J. J. Infrared Protein Crystallography. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics* **2011**, *1814* (6), 760–777.

(31) Whitmore, L.; Wallace, B. A. Protein Secondary Structure Analyses from Circular Dichroism Spectroscopy: Methods and Reference Databases. *Biopolymers: Original Research on Biomolecules* **2008**, *89* (5), 392–400.

(32) Narang, D.; Lento, C.; J. Wilson, D. HDX-MS: An Analytical Tool to Capture Protein Motion in Action. *Biomedicines* **2020**, *8* (7), 224.

(33) Maveyraud, L.; Mourey, L. Protein X-Ray Crystallography and Drug Discovery. *Molecules* **2020**, *25* (5), 1030.

(34) Billeter, M.; Wagner, G.; Wüthrich, K. Solution NMR Structure Determination of Proteins Revisited. *J Biomol NMR* **2008**, *42* (3), 155–158.

(35) Zhao, X. Protein Structure Determination by Solid-State NMR. *NMR of Proteins and Small Biomolecules* **2012**, 187–213.

(36) Yip, K. M.; Fischer, N.; Paknia, E.; Chari, A.; Stark, H. Atomic-Resolution Protein Structure Determination by Cryo-EM. *Nature* **2020**, *587* (7832), 157–161.

(37) Berkowitz, S. A.; Engen, J. R.; Mazzeo, J. R.; Jones, G. B. Analytical Tools for Characterizing Biopharmaceuticals and the Implications for Biosimilars. *Nat Rev Drug Discov* **2012**, *11* (7), 527–540. https://doi.org/10.1038/nrd3746.

(38) Alsenaidy, M. A.; Jain, N. K.; Kim, J. H.; Middaugh, C. R.; Volkin, D. B. Protein Comparability Assessments and Potential Applicability of High Throughput Biophysical Methods and Data Visualization Tools to Compare Physical Stability Profiles. *Front Pharmacol* **2014**, *5*, 39. https://doi.org/10.3389/fphar.2014.00039.

(39) Sörgel, F.; Lerch, H.; Lauber, T. Physicochemical and Biologic Comparability of a Biosimilar Granulocyte Colony-Stimulating Factor with Its Reference Product. *BioDrugs* **2010**, *24* (6), 347–357. https://doi.org/10.2165/11585100-000000000-00000.

(40) Zhou, D. H.; Shah, G.; Cormos, M.; Mullen, C.; Sandoz, D.; Rienstra, C. M. Proton-Detected Solid-State NMR Spectroscopy of Fully Protonated Proteins at 40 KHz Magic-Angle Spinning. *J Am Chem Soc* **2007**, *129* (38), 11791–11801.

(41) Zhou, D. H.; Shah, G.; Mullen, C.; Sandoz, D.; Rienstra, C. M. Proton-Detected Solid-State NMR Spectroscopy of Natural-Abundance Peptide and Protein Pharmaceuticals. *Angewandte Chemie* **2009**, *121* (7), 1279–1282.

(42) Linser, R.; Dasari, M.; Hiller, M.; Higman, V.; Fink, U.; Lopez del Amo, J.; Markovic, S.; Handel, L.; Kessler, B.; Schmieder, P. Proton-Detected Solid-State NMR Spectroscopy of Fibrillar and Membrane Proteins. *Angewandte Chemie International Edition* **2011**, *50* (19), 4508–4512.

(43) Liu, J.; Wu, X.; Zeng, Y.; Hu, Z.; Lu, J. Solid-State NMR Studies of Amyloids. *Structure* **2023**.

(44) Daskalov, A.; Martinez, D.; Coustou, V.; El Mammeri, N.; Berbon, M.; Andreas, L. B.; Bardiaux, B.; Stanek, J.; Noubhani, A.; Kauffmann, B. Structural and Molecular Basis of Cross-Seeding Barriers in Amyloids. *Proceedings of the National Academy of Sciences* **2021**, *118* (1), e2014085118.

(45) Bahri, S.; Silvers, R.; Michael, B.; Jaudzems, K.; Lalli, D.; Casano, G.; Ouari, O.; Lesage, A.; Pintacuda, G.; Linse, S. 1H Detection and Dynamic Nuclear Polarization–Enhanced NMR of Aβ1-42 Fibrils. *Proceedings of the National Academy of Sciences* **2022**, *119* (1), e2114413119.

(46) Shi, C.; Fricke, P.; Lin, L.; Chevelkov, V.; Wegstroth, M.; Giller, K.; Becker, S.; Thanbichler, M.; Lange, A. Atomic-Resolution Structure of Cytoskeletal Bactofilin by Solid-State NMR. *Sci Adv* **2015**, *1* (11), e1501087.

(47) Sprangers, R.; Velyvis, A.; Kay, L. E. Solution NMR of Supramolecular Complexes: Providing New Insights into Function. *Nat Methods* **2007**, *4* (9), 697–703.

(48) Asami, S.; Reif, B. Accessing Methyl Groups in Proteins via 1H-Detected MAS Solid-State NMR Spectroscopy Employing Random Protonation. *Sci Rep* **2019**, *9* (1), 1–13.

(49) Majumder, S.; Bhattacharya, D. S.; Langford, A.; Ignatius, A. A. Utility of High Resolution 2D NMR Fingerprinting in Assessing Viscosity of Therapeutic Monoclonal Antibodies. *Pharm Res* **2022**, *39* (3), 529–539.

(50) Arbogast, L. W.; Brinson, R. G.; Marino, J. P. Mapping Monoclonal Antibody Structure by 2D 13C NMR at Natural Abundance. *Anal Chem* **2015**, *87* (7), 3556–3561.

(51) Rosenzweig, R.; Kay, L. E. Bringing Dynamic Molecular Machines into Focus by Methyl-TROSY NMR. *Annu Rev Biochem* **2014**, *83*, 291–315.

(52) Purusottam, R. N.; Tekely, P. Site-Specific Dynamics of Methyl Groups Probed by Temporal Evolution of Heteronuclear 13C {1H} Overhauser Polarisation Encoded in 13C–13C Solid-State NMR Correlation Experiments. *Chem Phys Lett* **2020**, *754*, 137628.

(53) Wiesner, S.; Sprangers, R. Methyl Groups as NMR Probes for Biomolecular Interactions. *Curr Opin Struct Biol* **2015**, *35*, 60–67.

(54) Zinke, M.; Fricke, P.; Lange, S.; Zinn-Justin, S.; Lange, A. Protein− Protein Interfaces Probed by Methyl Labeling and Proton-Detected Solid-State NMR Spectroscopy. *ChemPhysChem* **2018**, *19* (19), 2457–2460.

(55) Nag, R.; Joshi, S.; Rathore, A. S.; Majumder, S. Profiling Enzyme Activity of L-Asparaginase II by NMR-Based Methyl Fingerprinting at Natural Abundance. *J Am Chem Soc* **2023**.

(56) Dey, A.; Mitra, D.; Rachineni, K.; Khatri, L. R.; Paithankar, H.; Vajpai, N.; Kumar, A. Mapping of Methyl Epitopes of a Peptide-Drug with Its Receptor by 2D STDD-Methyl TROSY NMR Spectroscopy. *ChemBioChem* **2022**, *23* (23), e202200489.

(57) Han, B.; Liu, Y.; Ginzinger, S. W.; Wishart, D. S. SHIFTX2: Significantly Improved Protein Chemical Shift Prediction. *J Biomol NMR* **2011**, *50* (1), 43–57.

(58) Metz, G.; Wu, X. L.; Smith, S. O. Ramped-Amplitude Cross Polarization in Magic-Angle-Spinning NMR. *J Magn Reson A* **1994**, *110* (2), 219–227.

(59) Weingarth, M.; Tekely, P.; Bodenhausen, G. Efficient Heteronuclear Decoupling by Quenching Rotary Resonance in Solid-State NMR. *Chem Phys Lett* **2008**, *466* (4–6), 247–251.

(60) Brinson, R. G.; Marino, J. P.; Delaglio, F.; Arbogast, L. W.; Evans, R. M.; Kearsley, A.; Gingras, G.; Ghasriani, H.; Aubin, Y.; Pierens, G. K. Enabling Adoption of 2D-NMR for the Higher Order Structure Assessment of Monoclonal Antibody Therapeutics. In *MAbs*; Taylor & Francis, 2019; Vol. 11, pp 94–105.

(61) Wang, D.; Park, J.; Patil, S. M.; Smith, C. J.; Leazer Jr, J. L.; Keire, D. A.; Chen, K. An NMR-Based Similarity Metric for Higher Order Structure Quality Assessment among US Marketed Insulin Therapeutics. *J Pharm Sci* **2020**, *109* (4), 1519–1528.

(62) Norrman, M.; Hubálek, F.; Schluckebier, G. Structural Characterization of Insulin NPH Formulations. *European Journal of Pharmaceutical Sciences* **2007**, *30* (5), 414–423.

(63) Bramham, J. E.; Podmore, A.; Davies, S. A.; Golovanov, A. P. Comprehensive Assessment of Protein and Excipient Stability in Biopharmaceutical Formulations Using 1H NMR Spectroscopy. *ACS Pharmacol Transl Sci* **2020**, *4* (1), 288–295.

(64) Li, M.; Reichert, P.; Narasimhan, C.; Sorman, B.; Xu, W.; Cote, A.; Su, Y. Investigating Crystalline Protein Suspension Formulations of Pembrolizumab from MAS NMR Spectroscopy. *Mol Pharm* **2022**, *19* (3), 936–952.

(65) Andrew, E. R.; Eades, R. G. A Nuclear Magnetic Resonance Investigation of Solid Cyclo Hexane. *Proc R Soc Lond A Math Phys Sci* **1953**, *216* (1126), 398–412.

(66) KovaL'Akova, M.; Fričová, O.; Hutníková, M.; Hronský, V.; Olčák, D. Dynamics of ^{1}H-^{13}C Cross Polarization in Nuclear Magnetic Resonance of Poly (3-Hydroxybutyrate). *Acta Phys Pol A* **2017**, *131* (4), 1162–1164.

(67) Fyfe, C. A.; Lewis, A. R.; Chézeau, J.-M. A Comparison of NMR Distance Determinations in the Solid State by Cross Polarization, REDOR, and TEDOR Techniques. *Can J Chem* **1999**, *77* (11), 1984–1993.

(68) Müller, L.; Kumar, A.; Baumann, T.; Ernst, R. R. Transient Oscillations in NMR Cross-Polarization Experiments in Solids. *Phys Rev Lett* **1974**, *32* (25), 1402.

(69) Ramamoorthy, A.; Wei, Y.; Lee, D.-K. PISEMA Solid-State NMR Spectroscopy. *Annu Rep NMR Spectrosc* **2004**, *52*, 1–52.

(70) Huster, D.; Schiller, J.; Arnold, K. Comparison of Collagen Dynamics in Articular Cartilage and Isolated Fibrils by Solid-state NMR Spectroscopy. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine* **2002**, *48* (4), 624–632.

(71) Knight, M. J.; Webber, A. L.; Pell, A. J.; Guerry, P.; Barbet-Massin, E.; Bertini, I.; Felli, I. C.; Gonnelli, L.; Pierattelli, R.; Emsley, L. Fast Resonance Assignment and Fold Determination of Human Superoxide Dismutase by High-Resolution Proton-Detected Solid-State MAS NMR Spectroscopy. *Angewandte Chemie* **2011**, *123* (49), 11901–11905.

(72) Zhou, D. H.; Rienstra, C. M. High-Performance Solvent Suppression for Proton Detected Solid-State NMR. *J Magn Reson* **2008**, *192* (1), 167–172. https://doi.org/10.1016/j.jmr.2008.01.012.

(73) Bishnoi, A.; Kumar, S.; Joshi, N. Wide-Angle X-Ray Diffraction (WXRD): Technique for Characterization of Nanomaterials and Polymer Nanocomposites. In *Microscopy methods in nanomaterials characterization*; Elsevier, 2017; pp 313–337.

(74) Bauer, C.; Freeman, R.; Frenkiel, T.; Keeler, J.; Shaka, A. J. Gaussian Pulses. *Journal of Magnetic Resonance (1969)* **1984**, *58* (3), 442–457.

(75) Caravatti, P.; Bodenhausen, G.; Ernst, R. R. Selective Pulse Experiments in High-Resolution Solid State NMR. *Journal of Magnetic Resonance (1969)* **1983**, *55* (1), 88–103.

(76) Freeman, R. Selective Excitation in High-Resolution NMR. *Chem Rev* **1991**, *91* (7), 1397–1412.

(77) Zhang, Z.; Su, Y.; Xiao, H.; Yang, J. Selective Nuclear Magnetic Resonance Method for Enhancing Long-Range Heteronuclear Correlations in Solids. *J Phys Chem Lett* **2022**, *13* (27), 6376–6382.

(78) Tekely, P.; Brondeau, J.; Elbayed, K.; Retournard, A.; Canet, D. A Simple Pulse Train, Using 90 Hard Pulses, for Selective Excitation in High-Resolution Solid-State NMR. *Journal of Magnetic Resonance (1969)* **1988**, *80* (3), 509–516.

(79) Veshtort, M.; Griffin, R. G. High-Performance Selective Excitation Pulses for Solid-and Liquid-State NMR Spectroscopy. *ChemPhysChem* **2004**, *5* (6), 834–850.

(80) Billeter, M. Non-Uniform Sampling in Biomolecular NMR. *Journal of biomolecular NMR*. Springer 2017, pp 65–66.

(81) Palmer, M. R.; Wenrich, B. R.; Stahlfeld, P.; Rovnyak, D. Performance Tuning Non-Uniform Sampling for Sensitivity Enhancement of Signal-Limited Biological NMR. *J Biomol NMR* **2014**, *58* (4), 303–314.

(82) Maciejewski, M. W.; Stern, A. S.; King, G. F.; Hoch, J. C. Nonuniform Sampling in Biomolecular NMR. In *Modern Magnetic Resonance*; Springer, 2008; pp 1305–1311.

(83) Billeter, M.; Orekhov, V. *Novel Sampling Approaches in Higher Dimensional NMR*; Springer Science & Business Media, 2012; Vol. 316.

(84) Maciejewski, M. W.; Mobli, M.; Schuyler, A. D.; Stern, A. S.; Hoch, J. C. Data Sampling in Multidimensional NMR: Fundamentals and Strategies. *Novel sampling approaches in higher dimensional NMR* **2011**, 49–77.

(85) Paramasivam, S.; Suiter, C. L.; Hou, G.; Sun, S.; Palmer, M.; Hoch, J. C.; Rovnyak, D.; Polenova, T. Enhanced Sensitivity by Nonuniform Sampling Enables Multidimensional MAS NMR Spectroscopy of Protein Assemblies. *J Phys Chem B* **2012**, *116* (25), 7416–7427.

(86) Porat, G.; Goldbourt, A. Assessment of Non-Uniform Sampling Schemes in Solid State NMR of Bacteriophage Viruses. *Isr J Chem* **2019**, *59* (11–12), 1027–1038.

(87) Burakova, E.; Vasa, S. K.; Klein, A.; Linser, R. Non-Uniform Sampling in Quantitative Assessment of Heterogeneous Solid-State NMR Line Shapes. *J Biomol NMR* **2020**, *74* (1), 71–82.

(88) Hyberts, S. G.; Milbradt, A. G.; Wagner, A. B.; Arthanari, H.; Wagner, G. Application of Iterative Soft Thresholding for Fast Reconstruction of NMR Data Non-Uniformly Sampled with Multidimensional Poisson Gap Scheduling. *J Biomol NMR* **2012**, *52* (4), 315–327.

(89) Bostock, M.; Nietlispach, D. Compressed Sensing: Reconstruction of Non-uniformly Sampled Multidimensional NMR Data. *Concepts in Magnetic Resonance Part A* **2017**, *46* (2), e21438.

(90) Le Guennec, A.; Dumez, J.; Giraudeau, P.; Caldarelli, S. Resolution-enhanced 2D NMR of Complex Mixtures by Non-uniform Sampling. *Magnetic Resonance in Chemistry* **2015**, *53* (11), 913–920.

(91) Zhang, R.; Nishiyama, Y.; Ramamoorthy, A. Exploiting Heterogeneous Time Scale of Dynamics to Enhance 2D HETCOR Solid-State NMR Sensitivity. *Journal of Magnetic Resonance* **2019**, *309*, 106615.

(92) Roy, M.; Brader, M. L.; Lee, R. W.; Kaarsholm, N. C.; Hansen, J. F.; Dunn, M. F. Spectroscopic Signatures of the T to R Conformational Transition in the Insulin Hexamer. *Journal of Biological Chemistry* **1989**, *264* (32), 19081–19085.

(93) Shneine, J.; Voswinkel, M.; Federwisch, M.; Wollmer, A. Enhancing the TR Transition of Insulin by Helix-Promoting Sequence Modifications at the N-Terminal B-Chain. **2000**.

(94) Wan, Z.; Huang, K.; Hu, S.-Q.; Whittaker, J.; Weiss, M. A. The Structure of a Mutant Insulin Uncouples Receptor Binding from Protein Allostery: An Electrostatic Block to the TR Transition. *Journal of Biological Chemistry* **2008**, *283* (30), 21198–21210.

(95) Derewenda, U.; Derewenda, Z.; Dodson, E. J.; Dodson, G. G.; Reynolds, C. D.; Smith, G. D.; Sparks, C.; Swenson, D. Phenol Stabilizes More Helix in a New Symmetrical Zinc Insulin Hexamer. *Nature* **1989**, *338* (6216), 594–596.

(96) Kaarsholm, N. C.; Ko, H. C.; Dunn, M. F. Comparison of Solution Structural Flexibility and Zinc Binding Domains for Insulin, Proinsulin, and Miniproinsulin. *Biochemistry* **1989**, *28* (10), 4427–4435.

(97) Smith, G. D.; Dodson, G. G. The Structure of a Rhombohedral R6 Insulin Hexamer That Binds Phenol. *Biopolymers: Original Research on Biomolecules* **1992**, *32* (4), 441–445.

(98) Choi, W. E.; Borchardt, D.; Kaarsholm, N. C.; Brzovic, P. S.; Dunn, M. F. Spectroscopic Evidence for Preexisting T-and R-state Insulin Hexamer Conformations. *Proteins: Structure, Function, and Bioinformatics* **1996**, *26* (4), 377–390.

(99) Jacoby, E.; KrüGer, P.; Karatas, Y.; Wollmer, A. Distinction of Structural Reorganisation and Ligand Binding in the [T ↔ R] Transition of Insulin on the Basis of Allosteric Models. *Biol Chem Hoppe Seyler* **1993**, *374* (7–12), 877–886. https://doi.org/10.1515/bchm3.1993.374.7-12.877.

(100) Maltesen, M. J.; Bjerregaard, S.; Hovgaard, L.; Havelund, S.; Van De Weert, M. Analysis of Insulin Allostery in Solution and Solid State with FTIR. *J Pharm Sci* **2009**, *98* (9), 3265–3277.

(101) Ferrari, D.; Diers, J. R.; Bocian, D. F.; Kaarsholm, N. C.; Dunn, M. F. Raman Signatures of Ligand Binding and Allosteric Conformation Change in Hexameric Insulin. *Biopolymers - Biospectroscopy Section* **2001**, *62* (5), 249–260. https://doi.org/10.1002/bip.1020.

(102) Smith, G. D.; Ciszak, E.; Magrum, L. A.; Pangborn, W. A.; Blessing, R. H. R6 Hexameric Insulin Complexed with M-Cresol or Resorcinol. *Acta Crystallogr D Biol Crystallogr* **2000**, *56* (12), 1541–1548.

(103) Smith, G. D.; Pangborn, W. A.; Blessing, R. H. The Structure of T6 Human Insulin at 1.0 Å Resolution. *Acta Crystallogr D Biol Crystallogr* **2003**, *59* (3), 474–482.

(104) Smith, G. D.; Pangborn, W. A.; Blessing, R. H. Phase Changes in T3R_3^{\rm F} Human Insulin: Temperature or Pressure Induced? *Acta Crystallogr D Biol Crystallogr* **2001**, *57* (8), 1091–1100.

(105) Nakagawa, S. H.; Zhao, M.; Hua, Q.; Hu, S.-Q.; Wan, Z.; Jia, W.; Weiss, M. A. Chiral Mutagenesis of Insulin. Foldability and Function Are Inversely Regulated by a Stereospecific Switch in the B Chain. *Biochemistry* **2005**, *44* (13), 4984–4999.

(106) Weiss, M. A. The Structure and Function of Insulin: Decoding the TR Transition. *Vitam Horm* **2009**, *80*, 33–49.

(107) Vinther, T. N.; Norrman, M.; Ribel, U.; Huus, K.; Schlein, M.; Steensgaard, D. B.; Pedersen, T. Å.; Pettersson, I.; Ludvigsen, S.; Kjeldsen, T. Insulin Analog with Additional Disulfide Bond Has Increased Stability and Preserved Activity. *Protein Science* **2013**, *22* (3), 296–305.

(108) Antoszewski, A.; Lorpaiboon, C.; Strahan, J.; Dinner, A. R. Kinetics of Phenol Escape from the Insulin R6 Hexamer. *J Phys Chem B* **2021**, *125* (42), 11637–11649.

(109) Hassiepen, U.; Federwisch, M.; Mülders, T.; Wollmer, A. The Lifetime of Insulin Hexamers. *Biophys J* **1999**, *77* (3), 1638–1654.

(110) Berchtold, H.; Hilgenfeld, R. Binding of Phenol to R6 Insulin Hexamers. *Biopolymers - Peptide Science Section* **1999**, *51* (2), 165–172. https://doi.org/10.1002/(SICI)1097-0282(1999)51:2<165::AID-BIP6>3.0.CO;2-X.

(111) Ciszak, E.; Smith, G. D. Crystallographic Evidence for Dual Coordination around Zinc in the T3R3 Human Insulin Hexamer. *Biochemistry* **1994**, *33* (6), 1512–1517.

(112) Rahuel-Clermont, S.; French, C. A.; Kaarsholm, N. C.; Dunn, M. F. Mechanisms of Stabilization of the Insulin Hexamer through Allosteric Ligand Interactions. *Biochemistry* **1997**, *36* (19), 5837–5845.

(113) Vashisth, H.; Abrams, C. F. Ligand Escape Pathways and (Un) Binding Free Energy Calculations for the Hexameric Insulin-Phenol Complex. *Biophys J* **2008**, *95* (9), 4193–4204.

Soumya Ranjan Pujahari

Ph.D. Scholar, IIT Bombay.