Peer-reviewed version available at *Biomolecules* **2017**, *7*, 63; doi:10.3390/biom7030063

Review

In Quest for Improved Drugs against Diabetes: The Added Value of X-ray Powder Diffraction Methods

Fotini Karavassili, Alexandros Valmas, Stavroula Fili, Christos D. Georgiou and Irene Margiolaki *

Section of Genetics, Cell Biology and Development, Department of Biology, University of Patras, GR-26500 Patras, Greece; fkar@upatras.gr (F.K); filistavr@gmail.com (S.F); valmas@upatras.gr (A.V); c.georgiou@upatras.gr (C.D.G)

* Correspondence: imargiola@upatras.gr; Tel.: +30 2610997408

Abstract:

Human Insulin (HI) is a well-characterized natural hormone which regulates glycose levels into the blood-stream and is widely used for diabetes treatment. Numerous studies have manifested that despite significant efforts devoted to structural characterization of this molecule and its complexes with organic compounds (ligands), there is still a rich diagram of phase transitions and novel crystalline forms to be discovered. Towards the improvement of drug delivery, identification of new insulin polymorphs from polycrystalline samples, simulating the commercially available drugs, is feasible today via macromolecular X- ray powder diffraction (XRPD). This approach has been developed and is considered as a respectable method, which can be employed in biosciences for various purposes such as observing phase transitions and characterizing bulk pharmaceuticals. An overview of structural studies on human insulin complexes performed over the past decade employing both synchrotron and laboratory sources for XRPD measurements**,** is reported herein. This review aims to assemble all recent advances in diabetes treatment field in terms of drug formulation, verifying in parallel the efficiency and applicability of protein XRPD for quick and accurate preliminary structural characterization in large scale.

Keywords: insulin; phenolic derivatives; crystallography; powder diffraction

1. Introduction

Diabetes mellitus (DM), was one of the first diseases ever described [1], whereas its name, was originated from the Greek word “Diabaino”(= passing through referring to the great emptying of the urine) and “meli”(= honey referring to the sweet taste of the patients urine due to high glucose concentration). Patients suffering from this disease (type I or II) lack control of glucose metabolism, due to inadequate levels and/or function of HI. Unfortunately, until Banting and Best’s isolation of insulin-containing extracts in 1926 [2], the prognosis for a patient was no better than it was over 3 millennials ago. Until now significant effort has been dedicated on the production of therapeutics for the control of DM and its treatment [3, 4], alleviating the daily routine for millions of patients. However diabetes, affecting nowadays a large and steadily increasing part of the world population [5, 6], still remains one of the most significant causes of morbidity and mortality in the world, and its global impact is likely to accelerate over the coming decades.

While type I DM [“insulin-dependent diabetes mellitus” or "juvenile diabetes”], which results from the pancreas's failure to produce enough insulin, reduces life expectancy of patients to almost 13 years, type II DM (related to failure of cells to respond properly to insulin) is not a directly life-threatening disease. Patients, however, experience a number of malfunctions of their circulatory system (hypertension, abnormal levels of cholesterol, triglycerides and blood sugars) as well as obesity [7,8] conditions which frequently induce numerous implications in human‘s health, such as nephropathy, neuropathy and several cardiovascular diseases (CVDs), including coronary heart disease (CHD), stroke and cardiomyopathy [9]. Recently, free radicals as well have been reported to

intervene in the onset of diabetes [10], while radical scavengers (such as salicylates) appear to effectively prevent diabetes and reducing severity of diabetic complications and especially myocardial infractions [11, 12].

Although there have been many improvements in diabetes' therapy it still remains a demanding challenge. Today, fifty years after HI's primary structural characterization [13], millions of patients rely on external provision of this essential pancreatic hormone of 5808 Da for their survival. Hypodermic injections, initially employed 80 years ago, containing microcrystals or balanced proportions of microcrystals and amorphous hormone, provide glycemic control over a certain period of time, depending on the timeframe of their action.

A *rapid*-acting insulin product, begins to act within 15 minutes after injection, reaches maximal levels in about 1 hour, and continues to act for 2 to 4 hours (e.g Insulin glulisine-Apidra, by Sanofi Aventis), while a *short*-acting one reaches the bloodstream within 30 minutes after injection and peaks from 2 to 3 hours later (e.g Novolin R, by Novo Nordisk) and a *long*-acting one reaches the bloodstream several hours after injection and tends to lower glucose levels fairly evenly over a 24-hour period (e.g Insulin detemir (Levemir), by Novo Nordisk). Slow dialysis of the injected clusters causes dissociation of insulin's hexameric structural motif (the typical storage form of crystalline insulin [14] into dimers and monomers, capable to diffuse through interstitial fluid, penetrate the capillary wall, and enter the blood circulation.

The aforementioned insulin microcrystals contain a mixture of polymorphs, distinguished either by the number of zinc ions per hexamer (two or four), or by the three distinct B chain conformations (extended-intermediate-helical) of the insulin monomer, denoted as "T", "R^f" and "R", naming the hexamers as T_6, $T_3R_3^f$ and R_6, respectively [15 - 17]. T_6 hexamers typically occur in the presence of zinc [18], while, addition of chloride and thiocynate anions results in $T_3R_3^f$ hexamers [19, 20]. Phenolic derivatives (phenol, resorcinol, m-cresol etc.), acting as ligands, enter inside corresponding pockets and strongly stabilize the hexameric form by hydrogen bonding the carbon oxygen of Cys^{A6} with the amide proton of Cys^{A11} [21], while in parallel assist towards the acquisition of the R_6 conformation, where residues B1 - B8 of each molecule adopt an a-helical configuration [22]. However, this transition is not always feasible [23, 24], while there is the unique, until now, case of the insulin analogue "ultralente" [14, 25], where addition of methylparaben, resulted in the reverse structural transition ($T_3R_3^f \rightarrow T_6$)[26]. The structural behavior of HI in the presence of several organic additives, mainly phenolic derivatives, which were originally used in pharmaceutical formulations as preservatives due to their strong antimicrobial activity, has been extensively examined [20, 28, 29].

B-chain conformational changes arise as well with respect to pH. This phenomenon is closely related to the charge ionizable groups of insulin adopt in distinct pH values. At pH values higher than 7.5, all carboxyl groups are negatively charged, one arginine and one lysine positively charged, all four Tyr and two His amino acids are neutral while the two amino termini are partially or largely unprotonated, depending on the exact pH [30].

Previous studies have established pH as an important factor affecting strongly protein crystallinity as well [31, 32]. It is known that protein solubility reaches a minimum near the isoelectric point (pI) whereas solubility increases in both directions away from it. Within a pH range – particular for each protein- molecules are altered in various ways. At certain pH values either Lys or Arg side chains start losing their positive charge or, in an alternative case, carboxyl groups of Asp and Glu side chains lose their negative charge [33]. Such partial amino acid neutralization disrupts the formation of salt bridges between protein molecules and thus decreases the crystallization rate. A lower degree of nucleation is likely to result in fewer but larger and better-formed crystals owing to the control of rapid crystal growth at low and high pH values.

Since the first crystallographic characterization of insulin, a number of polymorphs have been identified often related to distinguishing states of insulin [15, 19, 34]. The continuous study of human insulin's structural behavior (wild-type or various mutants) [20, 35] and the enormous amount of information extracted, allows for synchronization of crystalline insulin dissociation inside human body. Despite the resulting availability of various insulin products with rapid, intermediate

or prolonged action profiles, there is always a need for further improvement, as well as novel crystalline phases to be discovered [36]. Knowledge of the crystalline properties of individual polymorphs and their direct linkage with drug's ADME (Absorption, Distribution, Metabolism and Excretion) characteristics can lead to the production of more efficient formulations or even lead to the development of insulin preparations with alternative ways of delivery, such as sustained-release formulations or inhaled compounds [37].

As insulin products ideally contain millions of microcrystals, these should be studied as unity rather by isolating individual crystals to perform experiments, as single crystal X-ray diffraction (SCXD) approaches require. In contrary, powder diffraction methods can offer a holistic exploration of the system under investigation and can quickly provide accurate information while screening different polymorphs, advantages that clearly point out the value of this approach. Until the development of the Rietveld method [38], initially for neutron powder diffraction data and later for XRPD data [39], the most common practices of this method included phase identification and quantitative phase analysis, especially in industrial settings [40, 41]. However, powder diffraction currently holds a wider spectrum of applications including structure determination of zeolites, small organic molecules and most recently biological macromolecules [42, 43, 44].

The continuous evolution, during the last fifteen years, of macromolecular powder technique has considerably improved the quality of data which are routinely collected from microcrystalline precipitates using synchrotron or laboratory sources [45 - 47]. The simplicity of XRPD data collection and the uniqueness of the pattern obtained from each polymorph, suggest the powder diffraction technique as the most suitable approach for carrying out a quick and reliable characterization of various microcrystalline precipitates [48, 49]. This approach may provide medium resolution structural models (3 - 10 Å), however it offers the ability to study low quality crystals, polymorph screening becomes a routine practice and time-resolved studies are also possible. In addition, the size of the individual crystallites composing the powder can be directly estimated from the peak widths present in XRPD profiles. Further advantageous aspects deriving from the use of powder diffraction include homogeneity and purity control of the precipitates- valuable features for the pharmaceutical industry that cannot be obtained by SCXD measurements, as a single crystal may not be representative of the entire batch.

The main purpose of this review article is to summarize an extensive series of experiments and findings of the last decade while investigating the structural behavior of human insulin. Results include the structural characterization of HI in the presence or absence of small organic molecules (ligands) in a wide pH range (~4.5-8.5) obtained via XRPD, including the identification of four previously unknown insulin polymorphs; $C222_1$, $C2$, $P2_{1(\alpha)}$ & $P2_{1(\gamma)}$ [29, 50 - 53]. These studies indicate that there is still a rich diagram of human insulin phase transitions as well as novel crystalline phases to be discovered.

2. Overview of insulin crystalline structures

A broad selection of ligands was included in distinct crystallization experiments towards growing insulin crystal complexes, which were systematically studied under pH alterations. Polycrystalline precipitates (~50-100 μl volume each) of complexes, in the presence of zinc ions, were produced using the batch method, for almost all conditions studied.

Structural results, described in detail in the following sections, were obtained from diffraction patterns employing different instrumentation (Synchrotron sources: NSLS, Max-Lab, ESRF, SLS and laboratory: Rigaku RU200, PANalytical X'Pert PRO) in order to optimize data quality with respect to angular resolution (FWHM) and d-spacing resolution.

Prior to powder diffraction data collection, polycrystalline samples were loaded into borosilicate glass capillaries, sealed with grease to prevent dehydration, mounted on the diffractometer and spun to ensure adequate powder averaging. Synchrotron XRD experiments were carried out by applying capillary translation among scans in order to eliminate radiation damage effects (alterations in unit-cell parameters, peak broadening etc.) owing to the intense synchrotron

beam. Identical scans collected from newly exposed parts of each sample were summed together, in order to increase counting statistics without degrading data quality. In cases were laboratory instrumentation was employed, powder data didn't exhibit any radiation damage even after 24 hours of constant irradiation. The extraction of reliable lattice values and characterization of the peak shape and background coefficients was achieved, via the Pawley method [54].

2.1. First human insulin XRPD studies

Following the first successful experiment with polycrystalline metmyoglobin, conducted by R.B Von Dreele [44], which demonstrated that protein structure refinements using XRPD data are feasible, his research was further extended to insulin. Initially, microcrystalline slurry was produced as byproduct of a single-crystal sample [55] by grinding the crystals with mother liquor in an agate mortar [56]. The slurry was placed in a glass capillary and X-ray powder diffraction data were collected while the capillary was spun. Data collection was performed at room temperature at X3b1 beamline, at the National Synchrotron Light Source, equipped with a double Si (111) monochromator and a Ge (111) analyser.

From freshly made slurry, the diffraction pattern shown in Fig. 1(a) was obtained; whereas material left for 3 days after grinding produced a distinctly different diffraction pattern Fig. 1(b). The pattern from the grinded material was indexed in rhombohedral symmetry with a = 81.084, c = 37.537 Å, identical to the single-crystal unit cell for $T_3R_3^f$ HI conformation [19], whereas the pattern from the freshly ground material, revealed a previously unknown rhombohedral polymorph with a = 81.275, c = 73.024 Å, which is fundamentally a doubled c axis superlattice of the $T_3R_3^f$ structure (a phase denoted as $T_3R_3^f$DC).

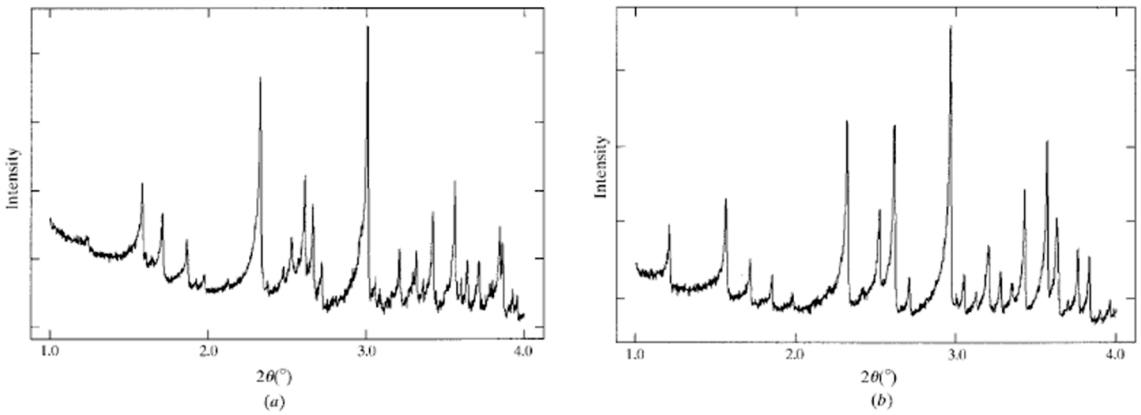


Figure 1. (a) X-ray powder diffraction pattern of freshly ground Zn-human insulin complex (T_3R_3DC) collected with λ = 0.700233 Å. (b) X-ray powder diffraction pattern of aged Zn-human insulin complex ($T_3R_3^f$) collected with λ = 0.700233 Å; the pattern shown was produced by the sum of two individual scans collected at 2s step^{-1} and 0.002° step^{-1} [Reproduction of Fig. 1 from reference 56].

Owing to the close relationship between these two phases, structure solution of $T_3R_3^f$DC using the molecular-replacement technique was employed. A starting model was introduced from the single-crystal coordinates for the $T_3R_3^f$ complex [19] and a three-parameter (two rotation angles and one translation) rigid-body Rietveld refinement was later performed. Atomic coordinates, extracted from stereochemically restrained Rietveld refinement of the $T_3R_3^f$ crystal structure, were used to complete the rigid-body refinement of the $T_3R_3^f$DC.

The complete structural characterization of the $T_3R_3^f$DC insulin form achieved via XRPD was also verified via single crystal experiments one year later [55] and revealed a number of special features of this new variant of the $T_3R_3^f$ human insulin-Zn complex. After grinding, a reduction of the material's volume by 2.095% or 1490 Å^{-3} per $T_3R_3^f$ complex was evident which consequently induced a structural change resulting in c axis doubling of the rhombohedral unit cell. One of the independent dimers rotates 17.2° about the c axis in the conversion from $T_3R_3^f$ to $T_3R_3^f$ DC; the other rotates 9.5° in the same direction. This rotation is accompanied by a collapse of the spacing between

pairs of $(AB)_2$ complexes along the crystallographic c axis and a repositioning of B chains with extended conformation. It is likely that water molecules extracted from the structure during grinding could come from this region.

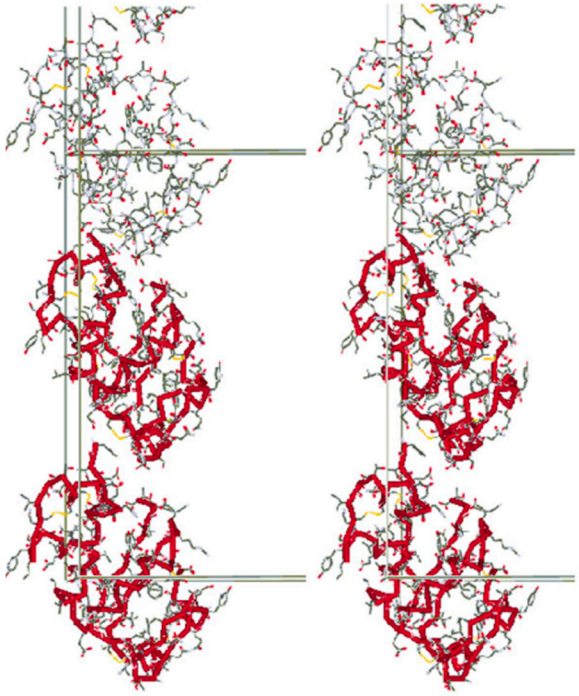


Figure 2. A stereographic packing diagram of three TR complexes stacked along the crystallographic c axis in $T_3R_3^fDC$. A Ca trace is shown in red for the two complexes that comprise the unique crystallographic unit and the unit-cell edges are also marked. [Reproduction of Fig. 4 from reference 56]

This was one of the first research results demonstrating the applicability of powder diffraction method for macromolecular crystal screening and detailed structure solution of a protein molecule. Within the next five years continuous developments in instrumentation as well as in data collection and analysis were carried out in parallel by Robert Von Dreele at Argonne National Laboratory (USA) and Irene Margiolaki and colleagues at the European Synchrotron Radiation Facility (ESRF, Grenoble, France). Their early studies on lysozyme (Turkey or Hen egg-white) as a model system, further established the use of XRPD as a valuable tool in the identification of small structural variations in protein molecules [46, 57 - 59].

2.2. Characterization of distinct insulin formulations via XRPD

Along with the underlying difficulties of developing and producing a biological formulation for therapeutic use, characterization of the final product can sometimes be even more challenging and demand a repeated revision process of analytical methods performed in a high throughput manner without compromising the accuracy of the obtained results. On top of this, protein therapeutics correspond to a class of products, which have a complex three dimensional structure whose integrity determines the bioavailability, biological activity, clinical efficacy and safety. All factors which control the aforementioned characteristics of a product are extensively studied in the production processes, and provide valuable information for further refining the enzyme/protein manufacturing.

The first study of this kind was originally conducted in 2006 by Norrman et al. [50], where 12 insulin formulations (some commercially available) were investigated via XRPD. Despite the medium-resolution XRPD patterns obtained, the data in combination with multivariate data analysis were used to compare insulin microcrystals preparations.

The commercially available insulin preparations examined in that project (Ultratard, Ultralente, Lente, Detemir, Penmix30, Novomix30 & Protaphan) were obtained from Novo Nordisk A/S, whereas additional microcrystals were prepared by batch crystallization. All products examined were "descendants" of the first stable protracted preparation of insulin, the NPH (Neutral Protamine Hagedorn) which was introduced in 1946 [60]. This formulation was based in an observation by Hans Christian Hagedorn (founder of former Novo Nordisk) and B. Norman Jensen in 1936, introducing that the effects of injected insulin could be prolonged with the addition of protamine –a peptide consisting mainly of arginine- obtained from the semen of river trout. An Insulin–zinc solution was cocrystallized with protamine, reducing insulin's solubility and resulting in NPH insulin; an intermediate–acting insulin product.

Among all HI crystals produced by batch crystallization, two novel crystal types were obtained. Orthorhombic $C222_1$ crystals (a = 59, b = 219, c = 223 Å) with three hexamers in the asymmetric unit, adopting the R_6 configuration were identified in presence of Urea, NaCl and resorcinol at pH 6.7 [50], whereas in slightly higher pH values (~7) monoclinic $C2$ crystals (a = 100 Å, b = 60 Å, c = 62 Å, β = 116°) were observed containing one hexamer with R_6 molecular conformation in the asymmetric unit and 50% solvent content [51]. Crystallization conditions for all formulations used in that study are summarized in Table 1.

Table 1. Crystallization conditions for the samples used in the study.

	A	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**	**J**	**K**	**X**
Space group	$P2_1$	$C222_1$	$I2_13$	$R3$	$R3$	$R3$	$R3$	$R3$	$P4_32_12$	$P4_32_12$	$P4_32_12$	$C2$
Insulin (mg ml^{-1})	5.2	3.5	10	14	3.8	3.8	3.8	1.5	3.8	3.8	1.5	3.5
Zn/hexamer	2.3	2.3		2.5	4	22	22	22	3	3	3	2.3
Phenol derivative (mM)†	20$^{(1)}$	25$^{(1)}$		19$^{(3)}$/19$^{(4)}$		65$^{(2)}$	65$^{(2)}$	65$^{(2)}$	7$^{(3)}$/14$^{(4)}$	7$^{(3)}$/14$^{(3)}$	7$^{(3)}$/14$^{(3)}$	25$^{(1)}$
NaCl (M)		1.0	1.0	0.02	0.3	0.12	0.12	0.12				1.0
Na acetate (M)					0.01	0.01	0.01	0.01				
Na citrate (M)				0.11								
Na_2HPO_4 (M)	0.48	0.05	0.04						0.013	0.013	0.013	0.05
Urea (M)		1.1										1.1
Tris (M)				0.14								
Protamine									Added in isophane ratio‡			
pH	7.3	6.7	7.2	8.15	5.5	7.4	7.4	7.4	7.3	7.3	7.3	7.0

† Phenol derivatives: (1) resorcinol (benzene-1,3-diol); (2) methyl p-hydroxybenzoate (methyl 4-hydroxybenzoate); (3) phenol; (4) m-cresol (3-methylphenol). ‡ Protamine added in isophane ratio: the protamine/insulin molar ratio that results in minimum protein concentration in the supernatant [60, 61]. [Reproduction of Table 2 from reference 50]

Protein powder data of this study were collected at room temperature, both in-house (rotating anode Rigaku RU200, Mar345 imaging plate, λ = 1.5418 Å) and at the Max-lab synchrotron (Lund, Sweden), beamlines 711, 911-2 and 911-3 [62, 63], using CCD detectors. Data indexing was in all cases unsuccessful, even though a variety of software wasexploited, due to low angular resolution (broad overlapping diffraction peaks) and the use of area detectors, which resulted in further peak overlap. Thus, only synchrotron powder diffraction patterns were employed for extracting preliminary structural information, due to their advantageous d-spacing and angular resolution. Nevertheless medium-resolution powder diffraction patterns were enough for effective classification in crystal systems via Principal Component Analysis (PCA) [64]. Crystallographic properties of all samples described in this project are listed in Table 2.

Table 2. Crystallographic properties of all insulin samples employed for structural investigation.

							UNIT CELL				
Crystal	**λ (Å)† / beamline**	**Trade name**	**Crystal system**	**Sp. group**	**Seq. origin‡**	**B-chain configuration**	***a*** **(Å)**	***b*** **(Å)**	***c*** **(Å)**	***β*** **(°)**	**PDB ref.§**
A	0.97 / 911-3		Monocl.	$P2_1$	H	R_6	61.3	61.7	47.5	111.3	1EV6[28]
B	1.00 / 911-2		Orthorh.	$C222_1$	H	R_6	58.9	219.4	223.7		In-house database
C	1.00 / 911-2		Cubic	$I2_13$	H	T	78.9	78.9	78.9		1APH[28]
D	0.97 / 911-3	Detemir	Rhomb	$R3$	H	R_6	78.9	78.9	39.5		1EV3[28]
E	0.97 / 911-3		Rhomb.	$R3$	H	$T_3R_3^f$	80.6	80.6	37.8		1TRZ[19]
F	0.969 / 711	Ultralente	Rhomb.	$R3$	H	T_6	81.3	81.3	33.7		1MSO[65]
G	0.969 / 711	Ultratard	Rhomb.	$R3$	H	T_6	82.5	82.5	34.0		4INS[18]
H	0.969 / 711	Lente			B, P						
I	0.969 / 711	Penmix30	Tetrag.	$P4_32_12$	H	R_6	62.9	62.9	85.9		In-house database
J	1.00 / 911-2	Novomix30	Tetrag.	$P4_32_12$	H B28Asp	R_6	62.8	62.8	86.9		In-house database
K	0.969 / 711	Protaphan	Tetrag.	$P4_32_12$	P	R_6	62.9	62.9	85.9		7INS[61]
X	0.97 / 911-2		Monocl.	$C2$	H	Unknown	100	60	62	116	

† Wavelength used during data collection. ‡ H = human, B = bovine, P = porcine. § Coordinate files used for simulated powder patterns [Reproduction of Table 1 from reference 50].

Patterns from different insulin polymorphs showed distinct peaks in the low-2θ region (0.9° to ~6°). Visual evaluation of the plots in Figure 3 shows that crystals, belonging to the same crystal system according to bibliography with same type of structure, have very similar powder patterns as well. Despite the fact that powder patterns have been collected without the optimum instrumentation, they reveal even small differences in protein structure based in alternations in peaks' positions (F, D and E crystals) and/or extinction of several peaks (I, J and K).

F, D and E crystals all belong to the rhombohedral space group $R3$ with T_6, R_6 and $T_3R_3^f$ molecular conformations, respectively. As seen from Figure 3 (left panel), similar peaks in the three patterns are generally shifted by less than 0.12° in 2θ. Peak variances are more evident within the 2θ range of 3.95 to 4.35°, where in all cases a high-intensity peak is observed, but its position is clearly different. The shifts in peak positions are associated to structural differences in the N-terminal part of the B-chain causing alternation in crystal packing and thus in the unit cell constants; especially in the length of the c-axis.

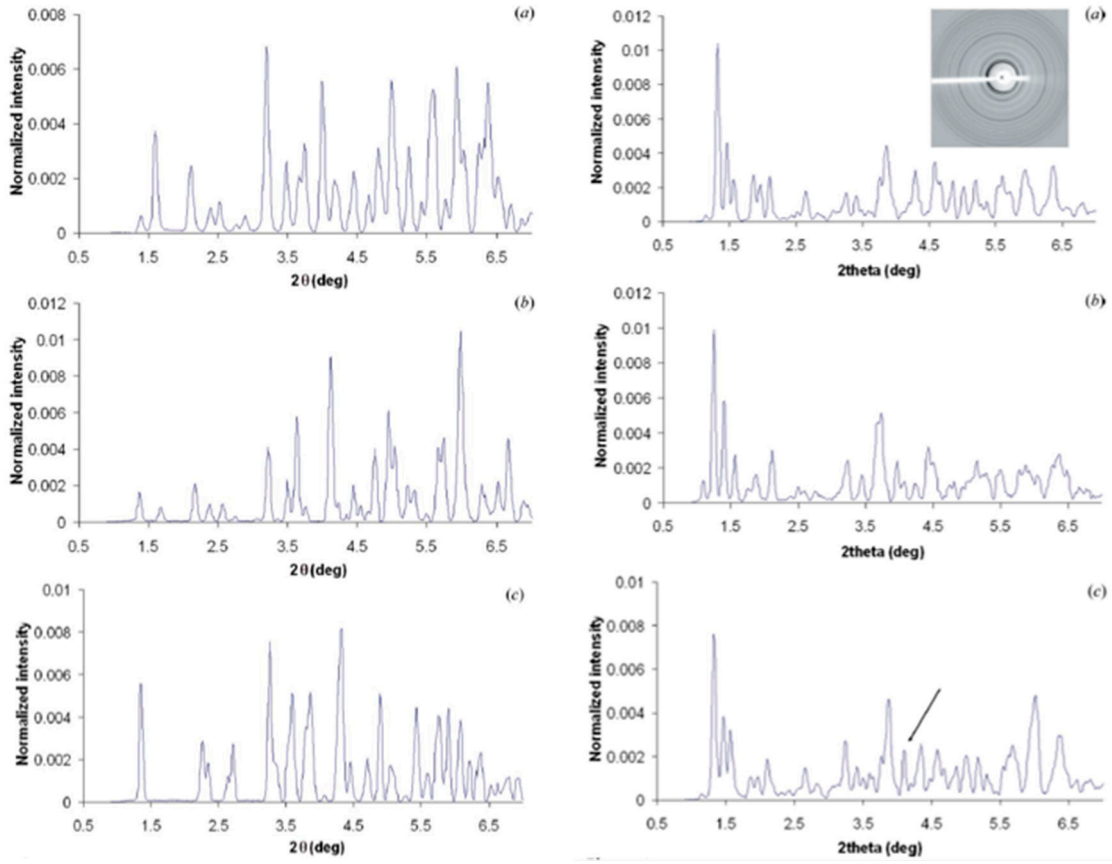


Figure 3. Left Panel: Normalized intensity versus 2θ for three types of rhombohedral crystals, D (a), E (b) and F (c). The three types correspond to distinct B-chain conformations (R_6, $T_3R_3^f$ and T_6, respectively). Differences are more evident between 3.95° and 4.35°, but there is a small shift in peak position throughout the whole 2θ region. Right Panel: Comparison of diffraction patterns from I (a), J (b) and K (c) crystals. I and K crystals have high degree of similarity whereas a slight shift in 2θ for the J crystals is evident, probably due to slightly larger unit-cell constants (Table 1). Inset is the diffraction pattern of the I crystals, obtained on a CCD detector at the beamline 711 (Max-lab, Sweden). The arrow in (c) indicates the extra peak found in the K crystals, but not in the I crystals. [Reproduction of Figure 2 & 3 from reference 50]

Crystals I, J and K belong to the same space group according to Table 2. Powder patterns from the three types of crystallites share a high degree of similarity especially in the low-2θ region, as shown in Fig. 3 (right panel). The major difference among them is an additional peak at 2θ = 4.1° in the K pattern (marked with an arrow in Fig. 3c) that is not found in the I pattern. Also, peak positions in J pattern are shifted relative to I and K patterns in the entire region reflecting the slightly larger unit-cell constants of J crystals (Table 1). This can be explained considering the mutation B28Asp in J crystals, which alters the molecule's charge, thus higher proportion of the protamine peptide is being bound on insulin [61] resulting in slightly altered unit-cell constants.

Visual analysis of the powder patterns described above, demonstrates that even without successful data indexing, the method can be used to distinguish different crystal systems and assess homogeneity of different batches or preparations of insulin. However, the complexity increases when examining a plethora of microcrystal suspensions and the procedure can be time-consuming, thus Norrman et al., employed the PCA analysis to facilitate the interpretation of powder patterns. Through PCA, data dimensionality (number of variables) is reduced, via a statistical procedure, from several hundreds to two or three principal components, resulting in a visual representation of the relationships and similarities of the –powder patterns of the- samples, by grouping them into clusters. Diffraction patterns from the above mentioned crystals were represented as data points and their clustering indicated a high similarity feature within each group. For example, the relative shifts in peak positions of the three rhombohedral D, E and F crystals, due to distinct B-chain conformations (R_6, $T_3R_3^f$ and T_6 respectively) had a large impact on the distribution of their PCA

scores in the plot and thus were not clustered together. Following this approach, different crystal systems and/or structural arrangements can be well separated, further facilitating the detection of novel polymorphs as in the case of B and X type of crystals which were clearly distinguished from other clusters.

The identification of two novel crystal forms (orthorhombic $C222_1$ and monoclinic $C2$) of human insulin accomplished in this project declare the use of XRPD as a powerful approach for characterization and evaluation of microcrystal suspensions of proteins, both during polymorph screening and in manufacturing process control. The medium-resolution data of the early XRPD era did not allow for detailed structural characterization, thus this was achieved a year later [51] via single crystal experiments (PDB codes: 2OM1 for the $C222_1$ crystal form and 2OLZ for the $C2$ one)

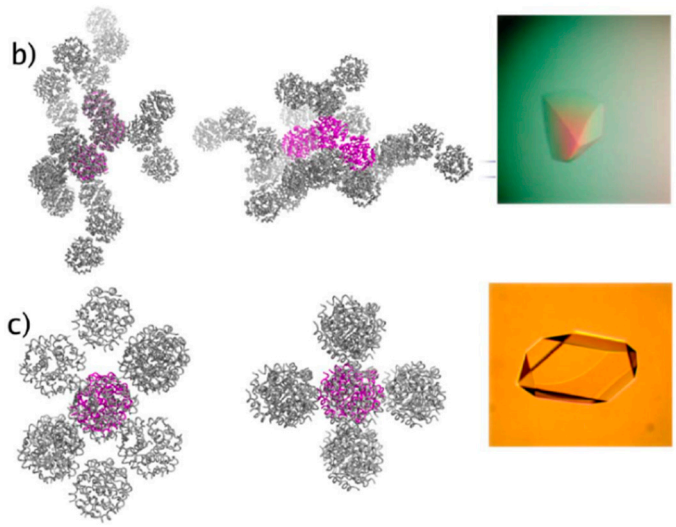


Figure 4. Crystal packing. (b) The crystal packing in the $C222_1$ space group drawn with main chain trace with the asymmetric unit in magenta. (c) The crystal packing of human insulin in $C2$ space group. The asymmetric unit molecule is colored magenta. The inserts in (b) and (c) show crystals of the $C222_1$ and $C2$ forms, respectively. [Reproduction of Figure 1 from reference 51]

The discovery of novel insulin polymorphs from Norrman & Schluckebier [51] triggered the research around insulin, and variations in cocrystallisation and pH conditions, forced the discovery of several other insulin crystalline polymorphs waiting to be examined in terms of physical stability, dissolution rate and other bioavailability properties.

Bovine insulin polycrystalline precipitates were extensively studied later on as well, in a wide pH range 5.0–7.6 [66]. Powder X-ray diffraction data revealed the T_6 hexameric insulin form (space group $R3$; unit-cell parameters $a = 82.59$, $c = 33.60$ Å for the sample crystallized at pH 5.0).

Fourteen powder diffraction profiles with slightly different lattice parameters, were selected for structure analysis. Lattice parameters variations were caused by alterations in sample preparation procedure or were induced by radiation exposure. These cause relative shifts in the positions of neighbouring peaks of diffraction patterns, allowing the contributing reflections of the overlapped peaks to be partially deconvoluted. Stereochemically restrained Rietveld refinement was performed to obtain an average crystal structure of bovine insulin over the pH range using the GSAS software [67, 68].

Selected regions of the refined coordinates and total OMIT map [69] computed at the final steps of analysis are presented in Figure 5. Each of the two zinc ions in the hexameric structure is octahedrally coordinated by three $N^{\varepsilon 2}$ atoms of three symmetry-related HisB10 residues and three symmetry related water molecules (PDB code: 4IDW).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 12 July 2017 doi:10.20944/preprints201707.0024.v1

Peer-reviewed version available at Biomolecules 2017, 7, 63; doi:10.3390/biom7030063

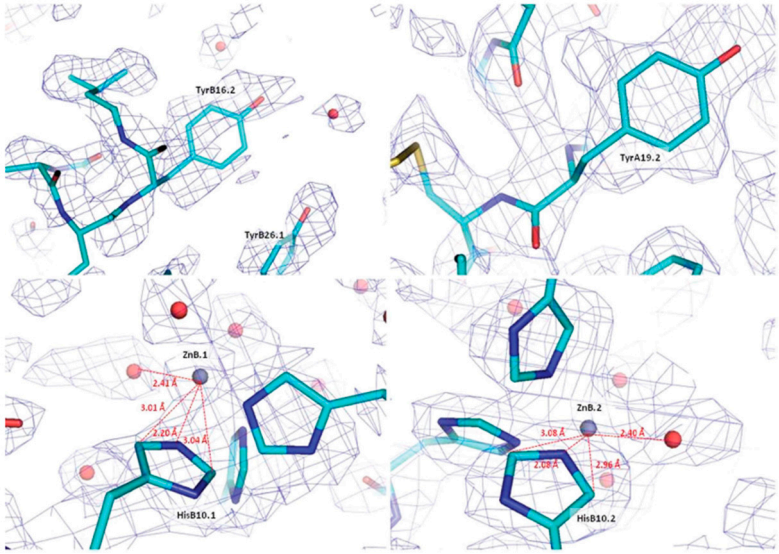


Figure 5. Selected regions of the final refined structural model in stick representation and the corresponding OMIT map contoured at 1. The cyan, blue and red colours in the stick representation illustrate C, N and O atoms of individual amino acids. The grey spheres represent the two independent Zn atoms and the red spheres denote water molecules identified during analysis. The closest distances between the two Zn ions and the neighbouring His atoms and water molecules are indicated in red. [Reproduction from reference 66]

The successful identification of the above formulations has reinforced the use of powder diffraction, by our group, as a rudimentary tool in daily research, for investigating the structural behavior of HI in a wide range of crystallization conditions in terms of pH and addition of ligands.

2.3. Cocrystallisation of HI with phenolic derivatives and pH dependence

Phenol and phenol-like compounds have been added in insulin formulations as antibacterial agents since the earliest years of production. It is well known that phenol binds in pockets of the insulin hexamer and has a dramatic effect on insulin's conformation, driving it to the R state [22].

While varying the pH in the presence of phenolic derivatives, a series of phase transitions has been reported. Specifically in the case of cocrystallization with phenol, four distinct polymorphs have been identified, three polymorphs with resorcinol, two with m-cresol and 4-nitrophenol and six with 4-ethylresorcinol (Table 3).

Table 3: Crystallographic properties of human insulin complexes with specific phenol derivatives (cell values reported, derive from an indicative sample within each pH range)

Phenol derivative	pH range	Space Group	Indicative sample's pH	*a* (Å)	*b* (Å)	*c* (Å)	*β* (°)	Resolution range (Å)
phenol								
	5.47-5.70	$P2_{1(\alpha)}$	5.70	114.682 (6)	337.63 (2)	49.270 (4)	101.555 (6)	112.2–7.5
	5.93-6.54	$C222_1$	6.14	60.287 (1)	221.797 (6)	228.812 (5)	90	115–7.5
	6.70-6.75	*C*2	6.75	103.0115 (5)	61.3213 (2)	63.5783 (4)	117.2244 (5)	45.9–5.3
	7.01-8.25	$P2_{1(\beta)}$	7.46	61.0920 (4)	61.8279 (4)	47.9302 (4)	110.6253 (7)	45–4.4
resorcinol								
	5.29-5.46	$P2_{1(\alpha)}$	5.29	114.0228 (8)	335.430 (3)	49.211 (6)	101.531 (8)	112.2–7.5
	5.93-7.45	*C*2221	6.40	60.5579 (7)	220.907 (3)	228.320 (3)	90	115–7.5
	7.53-8.22	$P2_{1(\beta)}$	8.22	61.0008 (4)	62.0040 (3)	47.8823 (3)	110.0465 (5)	45–4.4
m-cresol								
	4.50-6.70	$P2_{1(\gamma)}$		87.0749 (7)	70.1190 (5)	48.1679 (5)	106.7442 (8)	46.5-6.8

	6.70–8.60	*R*3 (R_6)	8.15	80.0644 (6)	80.0644 (6)	40.8396 (3)	90	40.5-3.7
4-nitrophenol								
	5.1–6.3	$P2_{1(\gamma)}$	5.97	87.118 (1)	70.9493 (9)	48.4967 (9)	106.653 (1)	46.5-6.8
	6.2 –8.1	*R*3 ($T_3R_3^f$)	6.41	80.721 (1)	80.721 (1)	37.8039 (5)	90	40.5-3.6
4-ethylresorcinol								
	4.95–5.60	$P2_{1(\gamma)}$	5.14	87.132 (3)	70.294 (2)	48.064 (2)	106.259(3)	47-6.5
	5.65-5.80	$P2_{1(\alpha)}$	5.80	114.130 (7)	336.086 (3),	48.987 (5)	101.935 (8)	112-12
	5.93–6.25	*C*2	5.97	103.0848(4)	61.6636 (2)	63.5006 (4)	117.417 (5)	46 - 7
	6.73–8.05	$P2_{1(\beta)}$	6.73	62.8231 (7)	62.1078 (5)	47.8362 (6)	111.6913 (9)	45 - 6

Quality of the obtained data allowed for successful indexing, using the fitted positions of at least 20 first reflections of each diffraction profile. From the extracted data, symmetry and unit-cell parameters were effectively determined.

When HI was crystallized with phenol, in addition to the earlier identified polymorphs $C222_1$, *C*2 [51] and $P2_1$ [28], a new monoclinic phase of insulin has been detected within the pH range 5.47–5.70, space group $P2_1$, [referred thereafter as $P2_{1(\alpha)}$]. Indexing of this unit cell was particularly challenging due to dominant-zone problem, where most of low-angle reflections belong to the dominant zone in reciprocal space. These reflections were not initially observed owing to peak overlap, however, combined use of diffraction data collected with different detectors confirmed the existence of a screw axis and led to the identification of the monoclinic cell $P2_{1(\alpha)}$ with remarkably large unit-cell parameters a = 114.682 (6), b = 337.63 (2), c = 49.270 (4) Å, β = 101.555 (6)°, which originally caused the dominant zone effect. Diffraction profiles acquired from $P2_{1(\alpha)}$ crystals extended to ~7.5 Å resolution. This was the first report of this specific crystallographic phase of human insulin.

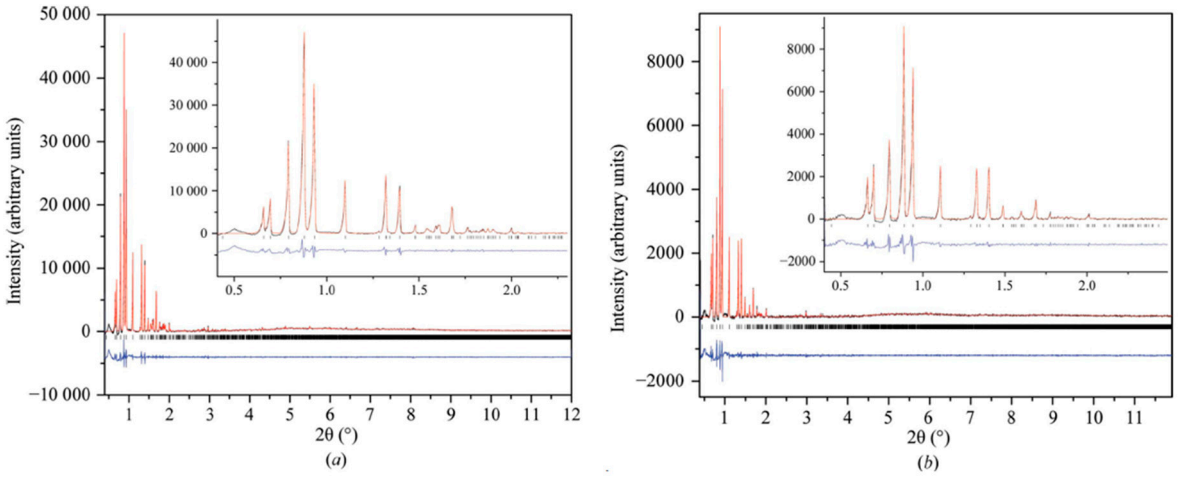


Figure 6. Pawley fit of the novel $P2_{1(\alpha)}$ phase. (a) Pawley fit of human insulin cocrystallized with phenol at pH 5.70. (b) Pawley fit of human insulin cocrystallized with resorcinol at pH 5.29. Data collection was performed at ESRF [λ = 1.299825 (16) Å, RT]. The black, red and lower blue lines represent the experimental data, the calculated pattern and the difference between the experimental and calculated profiles, respectively. The vertical bars correspond to Bragg reflections compatible with this particular space group. [Reproduction of Figure 6 from reference 29].

Crystallization of HI with resorcinol at pH 5.29 and 5.46 had essentially the same effect as with phenol and yielded the same monoclinic phase [space group $P2_1$, unit-cell parameters a = 114.0228 (8), b = 335.43 (3), c = 49.211 (6) Å, β = 101.531 (8)°].

The discovery of a previously unknown crystal form of insulin was the result of a systematic study of the effect of pH - even around its isoelectric point (~5.9) - on the crystallization behavior of

insulin in complex with zinc and a phenolic ligand. Nearby the pI region, its solubility is lowest and growing macroscopic crystals suitable for single-crystal X-ray structure determination is least likely to succeed. The novel insulin crystal packing, was identified in this exact pH area in the presence of phenol or resorcinol through powder diffraction and that is (probably) the reason why the monoclinic $P2_{1(\alpha)}$ conformation remained undetected even though crystallization experiments with phenol and resorcinol have been earlier reported [20, 27].

Nevertheless, the earlier identified insulin forms ($C222_1$ and $C2$) were obtained in these studies as well. Human insulin crystallized in the presence of phenol (pH 5.93–6.54) and resorcinol (pH 5.93–7.45) produced crystals with orthorhombic symmetry (space group $C222_1$) containing three protein hexamers per asymmetric unit [51].

In both cases, the pH increment resulted in minor alterations of unit-cell parameters illustrated by the smooth anisotropic shifts in the peak positions and no indication of a first-order phase transition. With the exception of the C2 phase, which was only observed during cocrystallization with phenol, all other phases obtained, coincided in crystallization experiments with the two ligands exhibiting minor alterations in unit-cell parameters.

Although phenol and resorcinol can substitute each other as allosteric ligands of the insulin hexamer without detectable changes in insulin structure [28], the choice of ligand apparently influences the crystallization behavior. This is noteworthy, as the phenolic binding site is buried and does not contribute to crystal contacts. The results show that human insulin crystallized in the presence of phenol and resorcinol can be greatly affected by pH. Systematic screening of crystallization conditions analysed by synchrotron XRPD yields an exact picture of insulin's crystallization behaviour. This analytical approach further extends the applicability of powder diffraction methods for efficient macromolecular crystal screening. The low instrumental contribution to the diffraction peaks and the high precision of determining unit-cell parameters allows small variations of the lattices to be quantified precisely.

The structural behavior of HI when cocrystallized with two widely used phenol-based ligands, *meta*-cresol (*m*-cresol) and 4-nitrophenol was further examined in a pH range [52]. These organic additives have been previously reported to be bactericidal agents and earlier structural results on HI complexed with these exist in the literature [27 - 29]. The use of *m*-cresol in particular has been suggested as a more effective germicide than phenol and is widely used as antimicrobial preservative in pharmaceutical formulations [70].

Several polycrystalline samples were produced and consecutive data collection experiments were performed using various X-ray sources to exploit their individual performances and to cross-check the results. A thorough data analysis revealed a first order phase transition with pH variation, resulting in two distinct polymorphs in both cases (Table 3), whereas a new monoclinic insulin phase has been found (space group $P2_1$, referred in the following as $P2_{1(\gamma)}$). Specifically when HI was crystallized with m-cresol (pH range 4.50–6.70) or 4-nitrophenol (pH range 5.1–6.3), this new monoclinic polymorph was identified (Fig. 7), with unit-cell parameters a = 87.0749 (7), b = 70.1190 (5), c = 48.1679 (5) Å, β = 106.7442 (8)°. Data acquired for the $P2_{1(\gamma)}$ polycrystalline samples extended to a d-spacing of approximately 6.8 Å.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 12 July 2017 doi:10.20944/preprints201707.0024.v1

Peer-reviewed version available at Biomolecules 2017, 7, 63; doi:10.3390/biom7030063

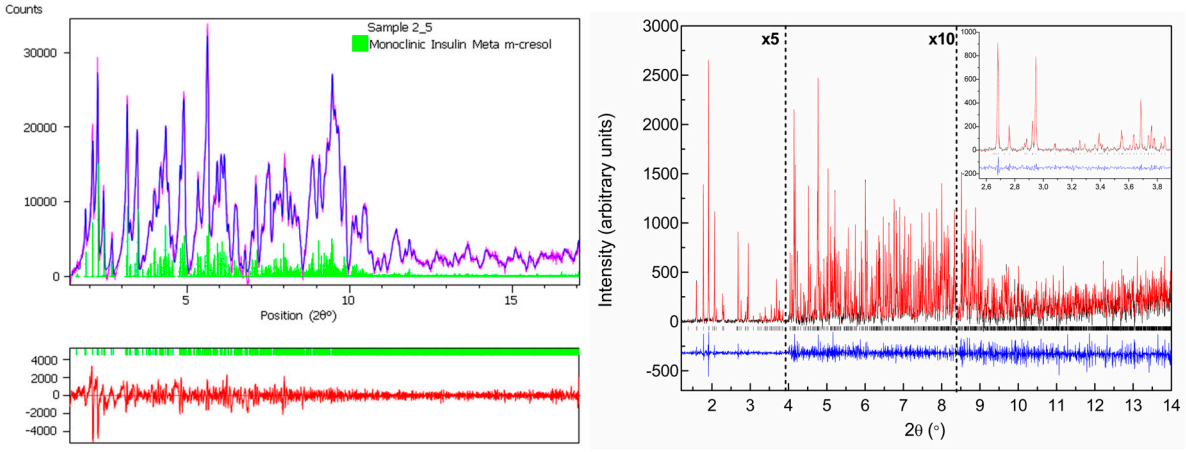


Figure 7. Pawley fit of the pattern of HI cocrystallized with m-cresol resulting in monoclinic crystal system, space group $P2_{1(\gamma)}$: (a) Data obtained from the X'Pert PRO laboratory diffractometer [λ=1.541874 Å, RT]. The pink, blue and lower red lines, represent the experimental data, the calculated pattern and the difference between the experimental and calculated profiles, respectively. The vertical green bars (lower left panel) correspond to Bragg reflections compatible with this particular space group ($P2_1$), while green bars in the upper panel correspond to the peak positions of the experimental pattern. (b) Data obtained from ID22 (former ID31) at ESRF [λ = 1.299944 (22) Å, RT]. The black, red and lower blue lines represent the experimental data, the calculated pattern and the difference between the experimental and calculated profiles, respectively. The vertical bars correspond to Bragg reflections compatible with the space group ($P2_1$). The profiles have been expanded by a factor of five at Bragg angles larger than 4° and by a factor of ten at angles larger than 8.5°. The inset corresponds to magnification of the fit in the selected 2θ range. [Reproduction of Figures 5 & 6 from reference 52]

Moving through the neutral pH region and reaching the weakly basic region, a first-order transition occurs, as it is evident in Figure 8. The monoclinic symmetry transforms quickly into a rhombohedral symmetry (space group $R3$) that is stable over a wide pH range (approximately 6.5–8.5) consisting of three protein hexamers per unit cell.

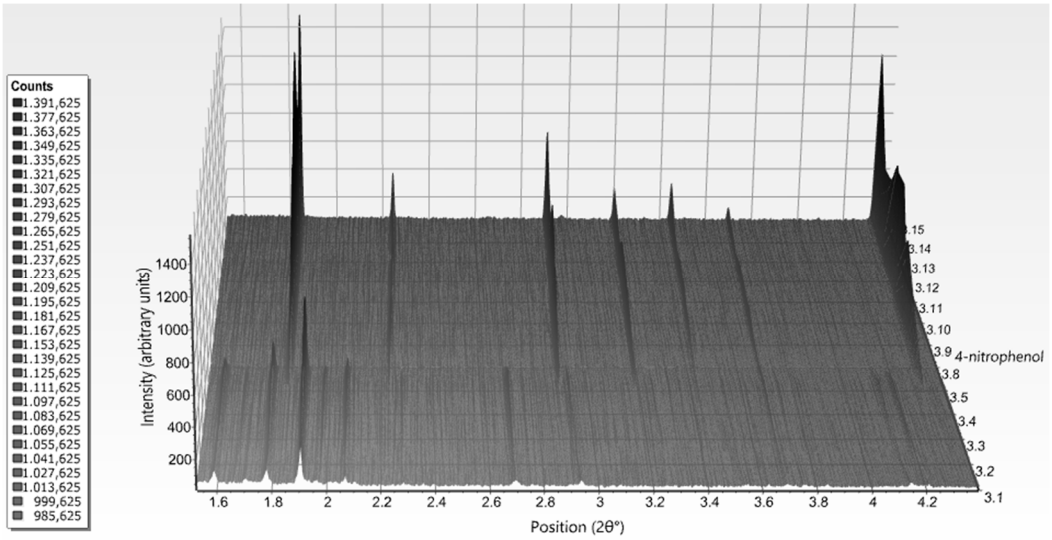


Figure 8. Overview of 15 diffraction patterns collected from crystals of HI in the presence of 4-nitrophenol, collected on ID22 (former ID31) [λ = 1.29989 (3) Å, RT]. The evolution of patterns within a pH range of 5.0- 8.5 is evident in this low 2θ range. [Reproduction of Figure 10, from reference 52]

Data analysis of powder diffraction profiles of HI complexed with 4-nitrophenol, led to the accurate extraction of the following unit-cell parameters a = 80.721 (1), c = 37.8039 (5) Å, γ = 120.000°

for the sample crystallized at pH 6.41. These values indicate that HI cocrystallized with this ligand acquires the $T_3R_3^f$ conformation [19]. The data collected on ID22 (former ID31) for these samples extended to a resolution of 3.6 Å.

When *m*-cresol is used as a ligand at pH 6.7–8.6, *R*3 space group is identified with slightly altered unit-cell parameters. Pawley analysis of high-resolution diffraction profiles, resulted in: a = 80.0644 (6), c = 40.8396 (3) Å, γ= 120.000° for the sample crystallized at pH 8.15. These values indicate that HI acquires the R_6 conformation [28]. Data collected on ID22 (former ID31) for these samples extended to a resolution of 3.7 Å. Consequently, the transition from the first crystal type to the second, causes a decrease in the unit-volume of about $\Delta V_{(P2_1(\gamma)\rightarrow R3)}/V_{P2_1(\gamma)}$ = -25.53%, while for HI complexed with m-cresol the decrease is almost 4% smaller ($\Delta V_{(P2_1(\gamma)\rightarrow R3)}/V_{P2_1(\gamma)}$ = -21.89%).

As opposed to the isosymmetrical polymorph, $P2_{1(\alpha)}$, identified by Karavassili *et al.* in 2012 [29], with surprisingly large a and b axes [a = 114.0228 (8), b = 335.43 (3), c = 49.211 (6) Å, β = 101.531 (8)°], the lattice parameters of this new polymorph $P2_{1(\gamma)}$ are remarkably smaller, approaching the already known range of dimensions that other known monoclinic cells adopt [22, 28]. Between these two monoclinic forms and the already deposited in the Protein Data Bank $P2_{1(\beta)}$ (PDB code: 1EV6; [28]), an unusual crystal packing for the $P2_{1(\gamma)}$ polymorph is noteworthy. The known $P2_{1(\beta)}$, consists of six molecules per asymmetric unit and 48% solvent content. According to Matthews Coefficient calculation [71, 72], the $P2_{1(\gamma)}$ polymorph contains twelve molecules per asymmetric unit and 39% solvent content (Matthews coefficient = 2.03 $Å^3$ Da^{-1}). This large difference reveals a much denser crystal packing in the case of $P2_{1(\gamma)}$ which could be of great pharmacological importance.

The additional interhexamer interactions that may occur owing to the very dense packing of the polymorph could associate with enhanced physicochemical properties and/or biological activity increasing the stability and thus extending the life of crystalline insulin formulations. Such an interpretation could possibly be a key point in the development of a new type of insulin-based microcrystalline compounds for diabetes treatment. Pharmaceutical products containing crystals with high protein concentration could lead minimize the injection times.

The complete structural model of the novel $P2_{1(\gamma)}$ polymorph has been derived by combining traditional single-crystal and emerging XRPD methods and will be presented in a forthcoming publication by our team [73].

The ligand 4-ethylresorcinol, a strong antiseptic and disinfectant of pharmaceutical formulations, was used during systematic crystallization experiments of HI in the presence of zinc ions as well [53]. Diffraction patterns obtained from several sources from crystals grown within the pH range 4.50–8.20 revealed four different crystalline polymorphs (Table 3). Among these, the two new monoclinic symmetry phases ($P2_{1(\alpha)}$ and $P2_{1(\gamma)}$) described earlier, were detected again, emphasizing their characterization as potential targets for the future development of microcrystalline insulin drugs.

The large quantity of diffraction patterns derived in this study were initially handled via PCA using *HighScore Plus* software [74] which classified patterns in four distinct groups (Fig. 9), corresponding to the mentioned crystalline phases, and indicated also the most representative sample of each cluster (marked with '***').

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 12 July 2017 doi:10.20944/preprints201707.0024.v1

Peer-reviewed version available at Biomolecules 2017, 7, 63; doi:10.3390/biom7030063

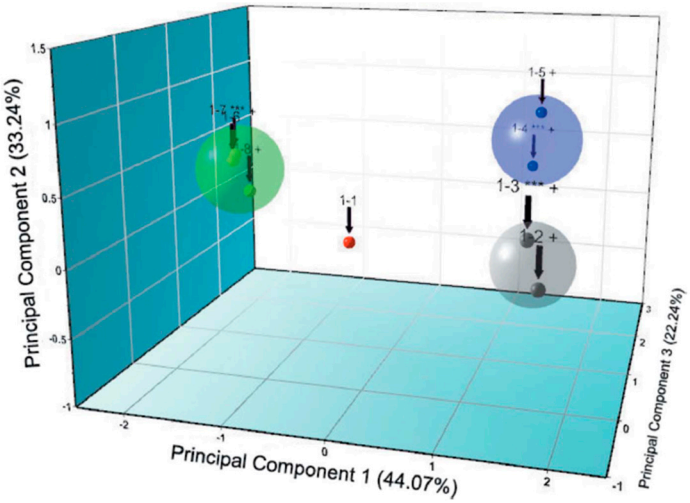
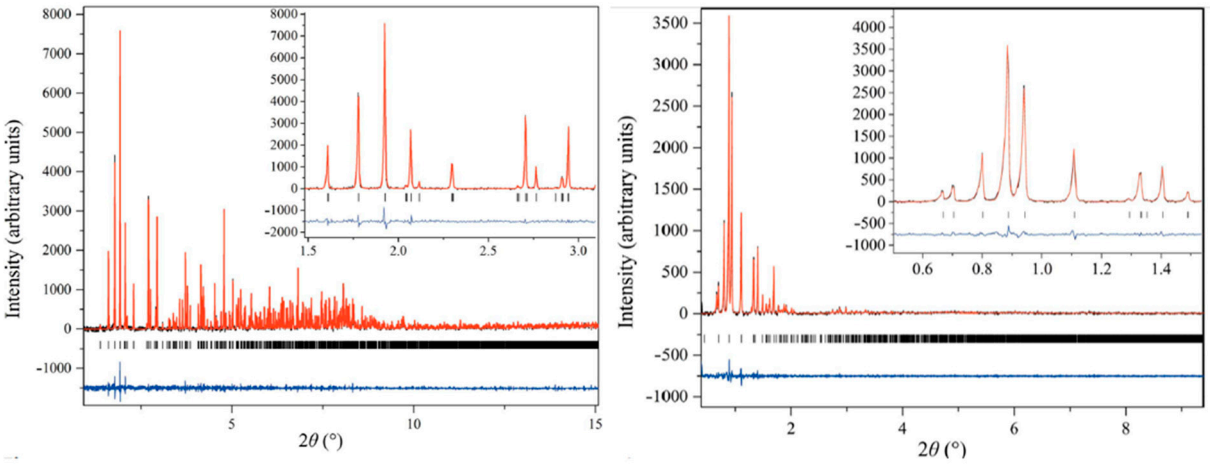


Figure 9. Cluster analysis of XRPD data of HI cocrystallized with 4-ethylresorcinol. Four distinct clusters were observed. Red cluster contains the pattern belonging to the new monoclinic symmetry $P2_{1(\gamma)}$, grey cluster contains patterns belonging to the $P2_{1(\alpha)}$ symmetry, blue cluster contains patterns belonging to the monoclinic symmetry $C2$ and green cluster contains patters belonging to $P2_{1(\beta)}$ symmetry. Numbers above each element correspond to sample coding. Data were collected on ID22 (former ID31) at the ESRF [λ = 1.29994 (1) Å, RT]. [Reproduction of Figure 4 from reference 53].

Systematic data analysis confirmed the three first order phase transitions with pH variation, observed in PCA analysis, which resulted in four distinct polymorphs of monoclinic symmetry (space group $P2_1$ and $C2$). Accurate unit-cell parameters of each polymorph are presented in Table 3.

Specifically, when HI was crystallized in the presence of 4-ethylresorcinol, within the pH range 4.95–5.80, two novel polymorphs with monoclinic symmetry [$P2_{1(\gamma)}$, in pH range *4.95–5.60* with unit-cell parameters a = 87.1323 (8), b = 70.294 (2), c = 48.064 (2) Å , β = 106.1729 (8)° and $P2_{1(\alpha)}$ in pH range 5.65-5.80 with unit-cell parameters a = 114.130 (7), b = 336.086 (3), c = 48.987 (5) Å , β = 101.935 (8)°] were observed. The same crystalline phases had been previously identified [29, 52] by XRPD. The data acquired for the $P2_{1(\gamma)}$ polycrystalline samples extended to a resolution of ~6.5 Å, whereas the lower resolution range for the $P2_{1(\alpha)}$ polycrystalline samples (~112-12 Å) was sufficient for successful indexing and Pawley analysis.



(a) (b)

Figure 10. Pawley fit of two diffraction patterns of HI cocrystallized with 4-ethylresorcinol collected on ID22 at ESRF [λ= 1.29994(1) Å, RT]: (a) At pH 5.02, polymorph $P2_{1(\gamma)}$. (b) At pH 5.80, polymorph $P2_{1(\alpha)}$. The black, red and lower blue lines represent the experimental data, the calculated pattern and the difference between the experimental and calculated profiles, respectively. The vertical bars correspond to Bragg reflections compatible with space group $P2_1$. [Reproduction of Figures 5 & from reference 53]

According to Matthews coefficient calculations [71, 72], the $P2_{1(\gamma)}$ phase contains 12 molecules (two hexamers) per asymmetric unit and 24 molecules (four hexamers) per unit cell, corresponding to a calculated solvent content of ~39.3% (Matthews coefficient of 2.03 $Å^3$ Da^{-1}). The unit-cell volume difference in the transition from $P2_{1(\gamma)}$ to $P2_{1(\alpha)}$ corresponds to an increase of about 6.5-fold. This is a large unit-cell modification. The $P2_{1(\alpha)}$ phase corresponds to one of the largest HI polymorphs ever been identified via XRPD methods to date, with the second belonging to the $C222_1$ symmetry {$V_{(C222_1)}$ = 3.054.394(63) $Å^3$[51], $V_{[P2_{1(\alpha)}]}$ = 1836620(73) $Å^3$ [29]}. As previously mentioned the complete structural model of the new $P2_{1(\gamma)}$ polymorph has been determined by combining traditional single-crystal and emerging analytical XRPD methods and will be presented in a forthcoming publication.

Finally, when HI was crystallized in the presence of 4-ethylresorcinol in the pH range ~6.00–8.00 the crystals obtained, belonged to monoclinic symmetry [space group $C2$ (pH 5.93–6.25) and $P2_{1(\beta)}$ (pH 6.73–8.05)]. The complete structural characterization of these polymorphs has been determined and thoroughly described previously (PDB code 2OLZ [51] and PDB code 1EVR [22, 28].

The systematic crystallization experiments of HI cocrystallized with the organic ligand 4-ethylresorcinol within the pH range 4.5–8.2 resulted in an accurate mapping of the symmetry and unit-cell parameters for all observed distinct monoclinic crystalline phases. Insulin in these polymorphs adopts the R_6 molecular conformation of B chain where binding interactions of ligands in the phenolic pockets seem to stabilize the specific conformation; a process assisted by a number of certain anions such as halides, pseudohalides and organic carboxylates.

Therefore, as most pharmaceutical preparations contain phenolic derivatives as preservatives, it is common HI molecules to adopt the $T_3R_3^f$ and R_6 conformations [75, 76]. As the stability level decreases from R to T, with R_6 being the most stable one [75], the existence of these conformations may function *in vivo* to create a balance between the stable storage and the gradual release of the active monomer. Moreover, the allosteric transition at the level of monomer could be important in terms of binding affinity between insulin and its receptor [77].

This could be of great importance in the case of injected treatments as it would be desirable to develop more effective formulations. Such improvement could be achieved by reducing crystal's dissolution rate and increasing the amount of active ingredient per dose. pH variation can result in distinct polymorphs with different physicochemical properties such as density, solubility and stability [78]. These characteristics can further affect their dissolution rate and thus their bioavailability. Therefore, the identification of new polymorphs could lead to the optimization of existing formulations or the design of advanced ones with a different action, depending on the needs of patients. It could also lead to the creation of new pharmaceutical preparations associated with alternative methods of administration, such as formulations with sustained release or formulations for inhaled administration [37]. Several HI polymorphs desrcibed in this study are summarized in Figure11 with respect to the ligand and the pH values each polymorph occurs.

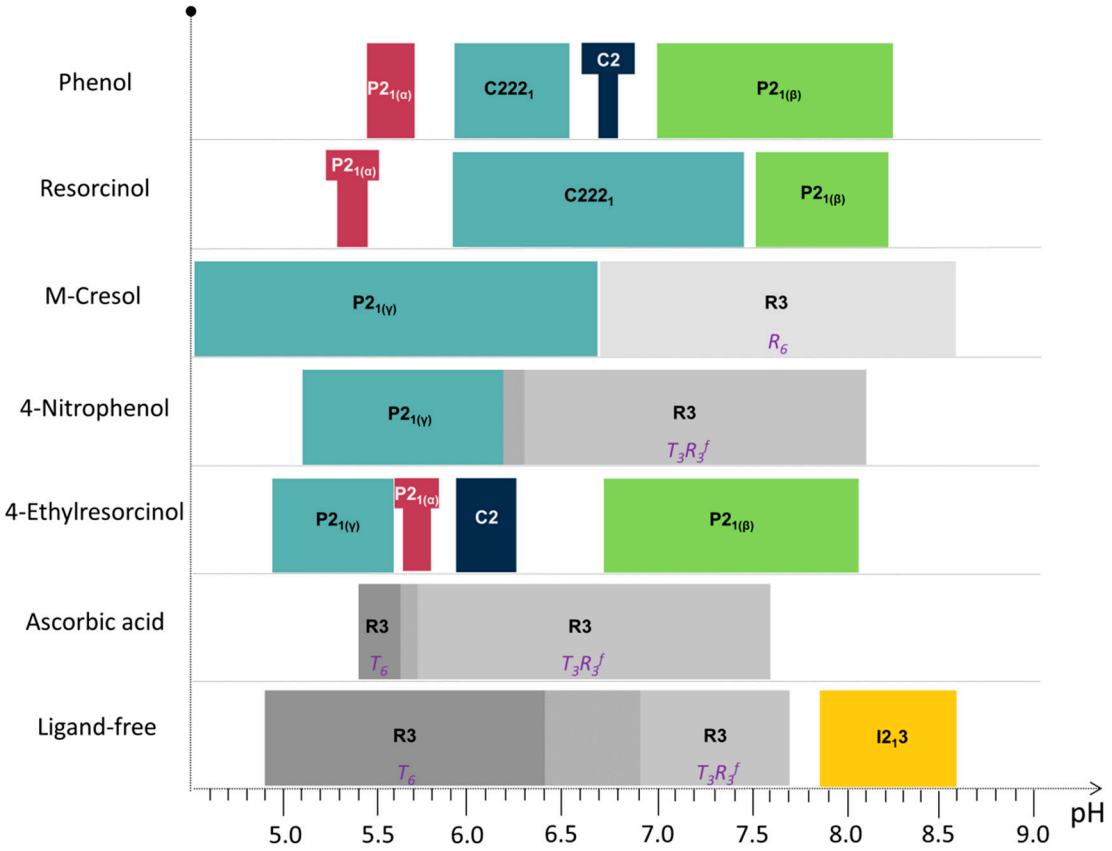


Figure 11. Phase diagram of HI polymorphs obtained through the described cocrystallizations while altering the pH. Dashed regions between polymorphs describe the simultaneous existence of both polymorphs. Molecular conformation obtained in rhombohedral symmetry is written in Italics. Exact pH values within each polymorph occurs are listed in Table 3.

2.4. Cocrystallisation of HI with a non-phenolic derivative and pH dependence

Increasing evidence in experimental and clinical studies suggest that oxidative stress plays a major role in the pathogenesis of both types of diabetes mellitus [79, 80] and causes complications affecting the vascular system, kidney, retina, lens, peripheral nerves, and skin [81].

Oxidative stress can be defined as the excess formation or/and insufficient removal of highly reactive molecules, such as reactive oxygen species (ROS) [82]. Free radicals are excessively produced in diabetics by glucose oxidation, nonenzymatic glycation of proteins, and the subsequent oxidative degradation of glycated proteins, all of which further overload the antioxidant system of patients. Thus, there is a necessity for introducing an overall treatment for controlling simultaneously insulin and antioxidants levels to minimize the diabetic's complications.

One of the strongest antioxidant substances [83], ascorbic acid (vitamin C) was selected for cocrystallisation with insulin, replacing the widely used phenol-based ligands (toxic in high concentrations). To date results indicate that HI has successfully cocrystallized with ascorbic acid in a pH range from 5.4 to 7.6. These new HI-ligand complexes could provide both insulin and free radical scavenger release over a certain period of time after entering blood stream. For example, the beneficial effects of ascorbic acid in diabetes mellitus and its health complications have been already demonstrated [84, 85].

Structural characterization of samples, in terms of unit-cell symmetry and dimensions, was performed via XRPD measurements employing both laboratory and synchrotron radiation. In the case where HI was cocrystallized in the pH range 5.40–5.65, diffraction patterns were typically indexed revealing the T_6 insulin conformation (space group: *R*3, a = 82.427(8) Å, c = 37.742(2) Å, for the sample crystallized at pH 5.44). Samples prepared in the pH range 5.70–7.66 were also found to

adopt the rhombohedral symmetry (space group: *R*3), however, extraction of unit-cell parameters indicates that insulin hexamers comprise the $T_3R_3^f$ conformation (a = 80.686(6) Å, c = 37.5868(1) Å, for the sample crystallized at pH 7.46). Patterns are shown in Figure 12.

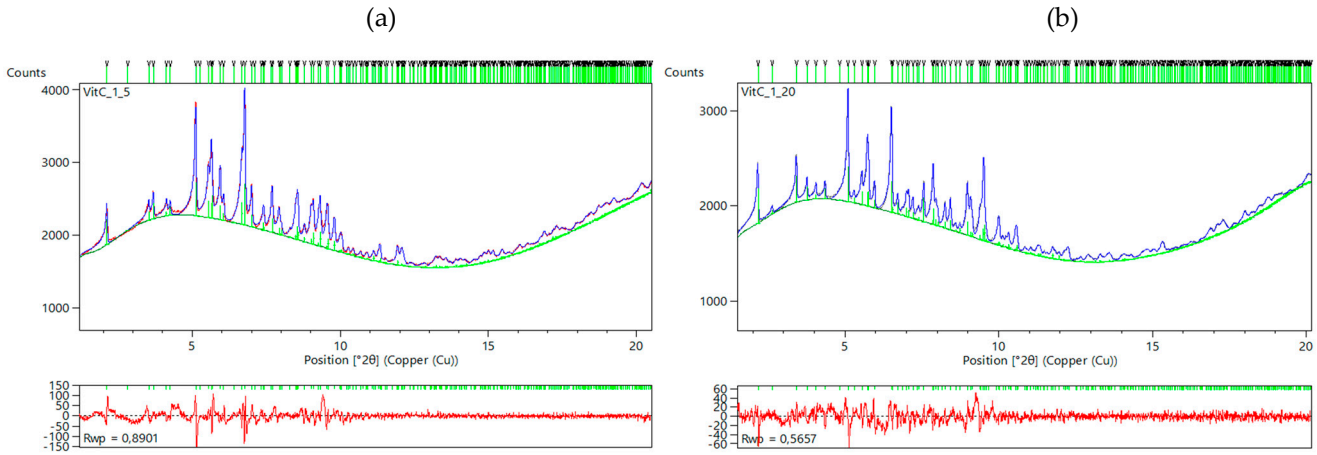


Figure 12. Diffraction patterns of HI cocrystallized with ascorbic acid. Data were collected using the X'Pert PRO laboratory diffractometer [λ=1.541874 Å, RT]. The red, blue and lower red lines represent the experimental data, the calculated pattern and the difference between the experimental and calculated profiles, respectively. The vertical green bars (lower panel) correspond to Bragg reflections compatible with this particular space group (*R*3), while green bars in the upper panel correspond to the peak positions of the pattern: (a) T_6 insulin conformation, (b) $T_3R_3^f$ insulin conformation.

2.5. Ligand-free crystalline HI studies and pH dependence

Towards the direction of understanding better the effect of pH upon HI conformational changes, further crystallization experiments were performed in a wide pH range (4.88-8.56) without the presence of any ligand.

Specifically, HI was crystallized using a solution of 13.14 mg/ml protein concentration, in the presence of 0.8 mM zinc acetate, 10.25 mM sodium thiocyanate and 0.4 M sodium/ monopotassium phosphate buffers of ascending pH per sample, in order to investigate the influence of pH on insulin crystallinity and conformation.

Diffraction patterns from both synchrotron and laboratory sources were collected and indexed successfully. Systematic data analysis led to the identification of 2 different space groups; *R*3 rhombohedral (pH range 4.9-7.7) and $I2_13$ cubic (pH range 7.75-8.60). Within the acidic pH, the T_6 configuration (space group: *R*3, a = 82.99 Å, c = 34.07 Å) was identified while in pH range 6.9 to 7.7 the T_6 alters to $T_3R_3^f$ (a = 80.66 Å, c = 37.74 Å), a transition that is evidently depicted in peak's positions changes (Fig. 13). The figure indicates an additional structural modification from samples with even higher pH values (7.8-8.6). A first order phase transition occurs at pH around 7.7 and insulin molecules obtain a cubic symmetry (space group: $I2_13$, a = 79.1 Å, PDB code: 9INS) [30].

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 12 July 2017 doi:10.20944/preprints201707.0024.v1

Peer-reviewed version available at Biomolecules 2017, 7, 63; doi:10.3390/biom7030063

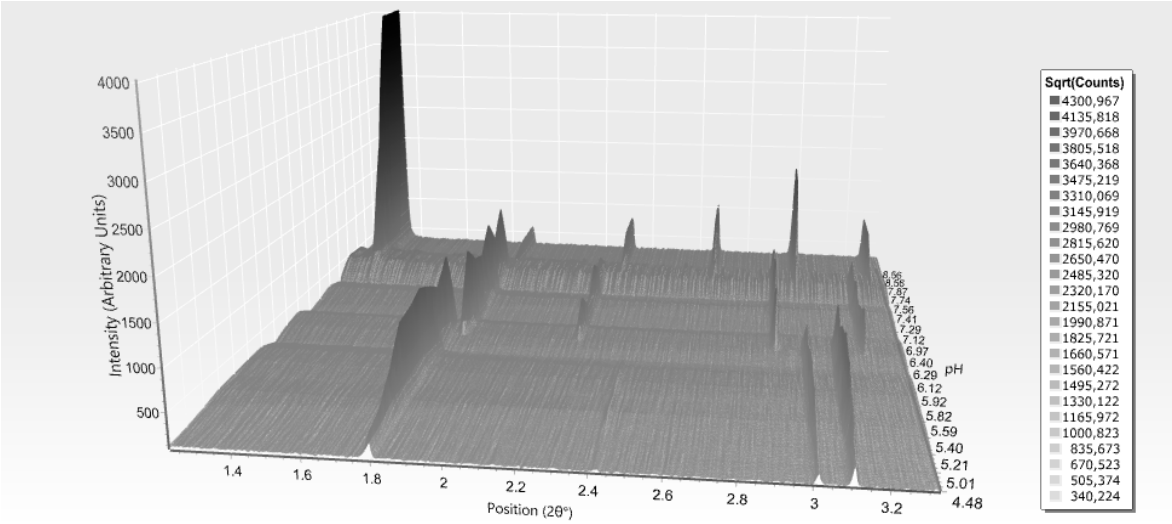


Figure 13. Surface plot showing the evolution of diffraction patterns from HI polycrystalline samples in the absence of ligands, at low 2θ range, while varying the pH (5.00-8.60). Patterns were collected on ID22 beamline at ESRF [λ=1.29974(6) Å, 280 K]. [86]

All diffraction patterns of this study were collected on ID22 at the ESRF and extend to a resolution of 3.3 Å (*R*3 polymorphs) and 2.7 Å (*I*2$_1$3 polymorph) as illustrated in Figure 15.

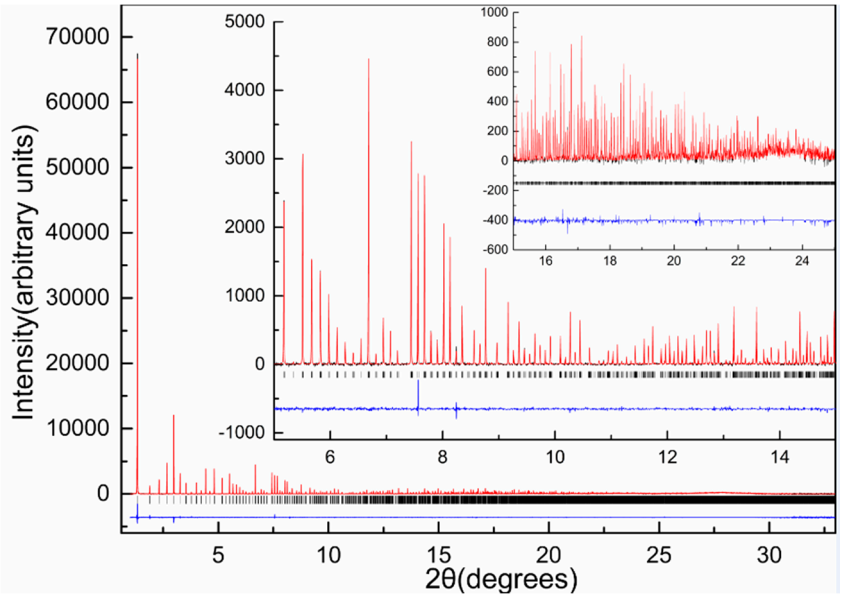
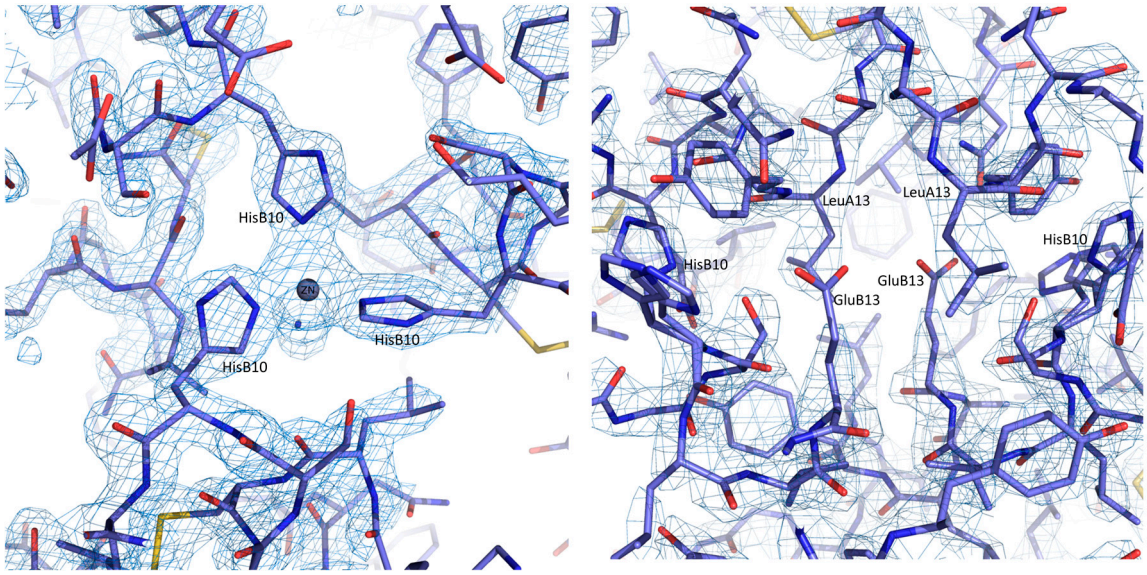


Figure 14. Pawley fit of diffraction pattern of HI crystallized in pH: 8.56 (space group: I2$_1$3), collected on ID22 at ESRF [λ= 1.29974(6) Å, 280 K]. The black, red and blue lines represent the experimental data, the calculated pattern and the difference between the experimental and calculated profiles, respectively. The vertical bars correspond to Bragg reflections compatible with space group I2$_1$3. Insets correspond to a magnification of the fit in the selected 2θ range. [87]

HI crystals grown in solutions with pH higher than 7.6 adopt the cubic symmetry, which is the most common zinc-free crystal form, in accordance with bibliography [88, 89]. However, zinc ions, mandatory for HI hexamer formation [90], were initially added during crystallisation. Consequently, we conclude that in alkaline conditions zinc ions are not able to interact with the molecule and for this reason HI crystals are formed from dimers and not from hexamers. The accuracy of this allegation was verified via structure solution of a microcrystalline sample at pH 8.56 from powder diffraction data (d-spacing resolution ~2.5 Å), which clearly revealed the absence of zinc ions from their common binding sites (Fig. 15): two identical high-affinity sites located on the

three-fold symmetry axis near histidines in the two distinct symmetries. Detailed description of the cubic structure will be discussed elsewhere [87].



(a) (b)

Figure 15. (a) Arrangement of three B10 histidines around a Zinc ion in rhombohedral symmetry crystals (space group *R*3, pH: 7.29). (b) Arrangement of B10 histidines in cubic symmetry crystals (space group $I2_13$, pH: 8.56), where the absence of zinc ions is evident. Histidines due to their flexibility are represented by 3 conformations [87].

This phenomenon can be explained considering the charge of all different ionizable groups of insulin molecule. For pH ≥ 7.5, histidines, due to imidazole rings' pKa (7.5), are neutral [91]. Uncharged His cannot associate with zinc ions and consequently insulin hexamers cannot form.

This observation could be of great importance for the pharmaceutical industry. The majority of the commercially available compounds consist of crystals containing HI hexamers, tightly packed within the unit-cell allowing minimum amount of solvent. However, it is evident that even slight alterations in storage conditions (e.g. temperature) which can directly affect parameters such as pH, may alter the tertiary molecular structure modifying physicochemical characteristics of the molecule and drug's ADME: Absorption, Distribution, Metabolism, and Excretion.

3. Discussion

The present review reports recent research advances of insulin-based polycrystalline compounds as potential therapeutics against diabetes. The majority of structural studies reported were conducted by employing macromolecular powder diffraction, a powerful complementary tool for swift and accurate structure determination of low quality crystalline material.

Constant improvements in protein engineering as well as the development of insulin analogs, introduced in the market, an important number of compounds capable to mimic up to a certain level the physiological secretion of insulin, accommodating both basal and prandial necessities. Empowered by the constant increase of diabetes cases among the population and the degenerative nature of the disease, insulin market has grown at a healthy 7% annual rate during the past decade [92]. This fact further motivates the worldwide drive for producing new insulin formulations and delivery systems in addition to the large portfolio of insulin products and analogs currently available from several manufacturers.

However new products need to be studied structurally to reveal specific characteristics (including homogeneity and purity) and polymorphism [93], defining physiological properties and clinical efficiency of formulations. This enormous need for structural data has consequently led to a

parallel development of techniques and approaches such as macromolecular powder diffraction for dynamic extraction of information even under challenging circumstances. In the last eighteen years, major advances have been made in the field of XRPD in terms of experimental methods and computational tools strengthening this technique and significantly expanding the variety of substances and samples that can be examined. XRPD data collection is simple, providing a distinct diffraction pattern for each polymorph within few minutes. That is exactly the reason the technique is pointed out as the most suitable tool for quickly and accurately characterizing numerous microcrystalline suspensions.

To date research findings on human insulin microcrystals exhibit a fascinating polymorphism, occurring upon physicochemical modifications of their environment (i.e. pH, ligand binding), further expanding the already rich phase diagram of the molecule. Four new biologically active types of HI crystals have been identified and their structures have been successfully determined by a combination of powder and single crystal diffraction measurements. Additionally, investigations were performed including cocrystallization of HI with a molecule of already proven pharmaceutical action, towards identifying the most beneficial complexes that will occur, which will lead to microcrystalline products of enhanced stability and activity.

Until the diabetes cure puzzle is completed, research for pharmaceutical products containing microcrystals with improved activity and stability will be at the center of scientific interest worldwide. Leading to a minimization of injection times, these compounds will be a life-quality improvement of great importance for millions of patients.

Acknowledgments: Authors are grateful to Professor A. N. Fitch, Dr J. P.Wright and Dr R. B. Von Dreele for their continuous input, advice and support during this research. We also thank Mathias Norrman and Gerd Schluckebier from Novo Nordisk, as well as PANalytical Company for supporting our macromolecular powder diffraction research. Part of the research conducted and presented in this review article has been funded by "RESEARCH PROJECTS FOR EXCELLENCE IKY/SIEMENS". Finally, we thank the NSLS, Max-Lab, ESRF, DESY and SLS synchrotrons for beam time allocation.

Conflicts of Interest: The authors declare no conflict of interest. The founding sponsors had no role in the design of the study; in the collection , analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results.

Abbreviations

The following abbreviations are used in this manuscript:

DM: Diabetes mellitus
HI: Human Insulin
CVD: Cardiovascular Disease
CHD: Coronary Heart Disease
ADME: Absorption, Distribution, Metabolism and Excretion
XRPD: X-Ray Powder Diffraction
SCXD: Sincle Crystal X-Ray Diffraction
ESRF: European Synchrotron Radiation Facility
PDB: Protein Data Bank
PCA: Principal Component Analysis

References

1. Ripoll, B.C.; Leutholtz, I. *Exercise and disease management*, 2nd ed.; CRC Press: Boca Raton, FL 2011; pp. 25, http://www.crcnetbase.com/doi/10.1201/b10856-5
2. Banting, F.G.; Best, C.H. The Internal secretion of the pancreas. *J. Lab. Clin. Med*. **1922**, *7*, 251–266.
3. Burn, P. Type 1 diabetes**.** *Nat. Rev. Drug. Discov.* **2010,** *9*, 187-188.

4. Waldron-Lynch, F.; Herold, K.C. Immunomodulatory therapy to preserve pancreatic β-cell function in type 1 diabetes. *Nat. Rev. Drug. Discov.* **2011,** *10*, 439-452.
5. Danaei, G.; Finucane, M.M.; Lu, Y.; Singh. G.M.; Cowan, M.J.; Paciorek, C.J.; Lin, J.K.; Farzadfar, F.; Khang, Y.H.; Stevens, G.A.; Rao, M.; Ali, M.K.; Riley, L.M.; Robinson C.A.; Ezzati, M. National, regional, and global trends in fasting plasma glucose and diabetes prevalence since 1980: systematic analysis of health examination surveys and epidemiological studies with 370 country-years and 2·7 million participants. *Lancet* **2011**, *378*, 31–40.
6. Lustig, R.H.; Schmidt, L.A.; Brindis, C.D. Public health: The toxic truth about sugar. *Nature* **2012**, *482*, 27–29.
7. Wilson P.W.; D'Agostino R.B.; Levy D.; Belanger A.M.; Silbershatz H.; Kannel W.B. Prediction of coronary heart disease using risk factor categories. *Circulation* **1998**, *97*:1837–1847.
8. McGill H.C.Jr.; McMahan, C.A; Determinants of atherosclerosis in the young. Pathobiological Determinants of Atherosclerosis in Youth (PDAY) Research Group. *Am. J. Cardiol*. **1998**, 82, 30T–36T.
9. Grundy, S.M. ; Benjamin, I.J.; Burke, G.L.; Chait, A.; Eckel, R.H.; Howard, B.V.; Mitch, W.; Smith, S.C. Jr; Sowers, J.R.; *Circulation* **1999**, *100*:1134-1146.
10. Lipinski B. Pathophysiology of oxidative stress in diabetes mellitus. *J. Diabetes Complicat*. **2001**, 15, 203-210.
11. Williamson, R.T.; Lond M.D. On the treatment of glycosuria and diabetes mellitus with sodium salicylate. *Brit. Med. J*. **1901**, 1, 760–762.
12. Ghiselli, A.; Laurenti, O.; DeMattia, G.; Maiani, G.; Ferro-Luzzi, A. Salicylate hydroxylation as an early marker of in vivo oxidative stress in diabetic patients. *Free Radical Bio. Med*. **1992**, 13, 621–626.
13. Adams, M. J.; Blundell, T. L.; Dodson, E. J.; Dodson, G.G.; Vijayan, M.; Baker, E.N.; Harding, M.M.;Hodgkin, D.C.; Rimmer, B. & Sheat, S. Structure of rhombohedral 2-zinc insulin crystals. *Nature* **1969**, *224*, 491–495.
14. Schlichtkrull, J. Chemical and biological studies on insulin crystals and insulin zinc suspensions, Thesis, Copenhagen University, **1958**.
15. 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*, 4427–4435.
16. Ciszak, E.; Beals, J.M.; Frank, B.H.; Baker, J.C.; Carter, N.D.; Smith G.D., Role of C-terminal B-chain residues in insulin assembly: the structure of hexameric LysB28ProB29-human insulin. *Structure* **1995**, *3* 615–622.
17. Norrman, M. Insulin polymorphism. PhD thesis, Lund University, Sweden, **2007**.
18. Baker, E.N.; Blundell, T.L.; Cutfield, J.F.; Cutfield, S.M.; Dodson, E.J.; Dodson, G.G.; Hodgkin, D.C.; Hubbard, R.E.; Isaacs, N.W.; Reynolds, C.D.; Sakabe, K.; Sakabe, N. & Vijayan, N.M. The structure of 2Zn pig insulin crystals at 1.5 Å resolution. *Philos. Trans. R. Soc. London Ser. B*, **1988**, *319*, 369-456.
19. Ciszak, E. & Smith, G. D. Crystallographic Evidence for Dual Coordination around Zinc in the T_3R_3 Human Insulin Hexamer. *Biochemistry* **1994**, *33*, 1512–1517.
20. Whittingham, J.L.; Chaudhurri, S.; Dodson, E.J.; Moody, P.C.E.; Dodson, G.G. X-ray Crystallographic Studies on Hexameric Insulins in the Presence of Helix-Stabilizing Agents, Thiocyanate, Methylparaben and Phenol. *Biochemistry* **1995**, *34*, 15553-15563.
21. Dunn, M.F. Zinc-ligand interactions modulate assembly and stability of the insulin hexamer - a review. *Biometals* **2005**, *18*, 295–303.
22. 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*, 594–596.
23. Smith, G.D. & Ciszak, E. The structure of a complex of hexameric insulin and 4'-hydroxyacetanilide. *Proc. Natl Acad. Sci. USA* **1994**, *91*, 8851-8855.
24. Smith, G.D.; Ciszak, E. & Pangborn, W.A. A novel complex of a phenolic derivative with insulin: structural features related to the T-->R transition. *Protein Sci*. **1996**, *5*, 1502-1511.
25. Hallas-Møhler, K.; Petersen, K.; Schlichtkrull, J. Crystalline and amorphous insulin–zinc compounds with prolonged action. *Science* **1952**, 116, 394 – 398.
26. Wagner, A.; Diez, J.; Schulze-Briese, C. & Schluckebier, G. Crystal structure of Ultralente - A microcrystalline insulin suspension. *Proteins* **2009**, *74*, 1018–1027.

27. Whittingham, J.L.; Edwards, D.J.; Antson, A.A.; Clarkson, J.M. & Dodson, C.G. Interactions of Phenol and *m*-Cresol in the Insulin Hexamer, and Their Effect on the Association Properties of B28 Pro → Asp Insulin Analogues. *Biochemistry* **1998**, *37*, 11516–11523.
28. Smith, G.D.; Ciszak, E.; Magrum, L.A.; Pangborn, W.A. & Blessing, R.H. R_6 hexameric insulin complexed with *m*-cresol or resorcinol. *Acta Cryst. D* **2000**, *56*, 1541–1548.
29. Karavassili, F.; Giannopoulou, A.E.; Kotsiliti, E.; Knight, L.; Norrman, M.; Schluckebier, G.; Drube, L.; Fitch, A.N.; Wright, J.P. & Margiolaki, I. Structural studies of human insulin cocrystallized with phenol or resorcinol via powder diffraction. *Acta Cryst*. D **2012**, *68*, 1632–1641.
30. Gursky, O.; Badger, J.; Li, Y. & Caspar, D.L. Conformational changes in cubic insulin crystals in the pH range 7-11. *Biophys. J.* **1992**, *63*, 1210–1220.
31. McPherson, A. Crystallization of Proteins by Variation of pH or Temperature. *Methods Enzymol*. **1985**, *114*, 125–127.
32. Farr, R.; Perryman, A. & Samudzi, C. Re-clustering the database for crystallization of macromolecules. *J. Cryst. Growth* **1998**, *183*, 653–668.
33. McPherson, A. Increasing the size of microcrystals by fine sampling of pH limits. *J. Appl. Cryst*. **1995**, *28*, 362–365.
34. Smith, G.D.; Pangborn, W.A. & Blessing, R.H. The structure of T_6 bovine insulin. *Acta Cryst*. D **2005**, *61*, 1476–1482.
35. Hodgkin, D.C. Insulin molecules: the extent of our knowledge. *Pure Appl. Chem*. **1971**, *26*, 375–384.
36. Brange, J. The new era of biotech insulin analogues. *Diabetologia* **1997**, *40*, S48-53.
37. 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*, 301–317.
38. Rietveld, H.M. A profile refinement method for nuclear and magnetic structures. *J. Appl. Cryst*. **1969**, *2*, 65–71.
39. Young, R.A.; Mackie, P.E. & Von Dreele, R.B. Application of the pattern-fitting structure-refinement method of X-ray powder diffractometer patterns. *J. Appl. Cryst*. **1977**, *10*, 262-269.
40. Alexander, L.E. Forty years of quantitative diffraction analysis. Proceedings of Advances in X-ray Analysis; McMurdie H.F.; Barrett C.S.; Newkirk J.B. & Ruud, C.O. Eds. Plenum Press: New York. 1976, *20*, 1-13.
41. Klug, H.P. & Alexander, L.E. X-ray Diffraction Procedures for Polycrystalline and Amorphous Materials. 2nd ed. John Wiley: New York, **1974**, pp. 271-418.
42. Cheetham, A.K. & Taylor, J.C. Profile analysis of neutron powder diffraction data: its scope, limitations and applications in solid state chemistry. *J. Solid State Chem*. **1977**, *21*, 253-275.
43. Harris, K.D.M. & Tremayne, M. Crystal Structure Determination from Powder Diffraction Data. *Chem. Mater*. **1996**, *8*, 2554-2570.
44. Von Dreele, R.B. Combined Rietveld and stereochemical restraint refinement of a protein crystal structure. *J. Appl. Cryst*. **1999**, *32*, 1084-1089.
45. Von Dreele, R.B. Protein Crystal Structure Analysis from High-Resolution X-Ray Powder-Diffraction Data. *Methods Enzymol*. **2003**, 368, 254–267.
46. Margiolaki, I.; Wright, J.P.; Fitch, A.N.; Fox, G.C. & Von Dreele, R.B. Synchrotron X-Ray Powder Diffraction Study of Hexagonal Turkey Egg-White Lysozyme. *Acta Cryst*. D **2005**, *61*, 423–432.
47. Margiolaki, I.; Wright, J.P.; Wilmanns, M.; Fitch, A.N. & Pinotsis, N. Second SH3 Domain of Ponsin Solved from Powder Diffraction. *J. Am. Chem. Soc*. **2007**, *129*, 11865–11871.
48. Margiolaki, I. & Wright, J.P. Powder crystallography on macromolecules. *Acta Cryst. A* **2008**, *64*, 169–180.
49. Karavassili, F. & Margiolaki, I. Macromolecular Powder Diffraction: Ready for genuine biological problems. *Protein Pept. Lett*. **2016**, *23*, 232–241.
50. Norrman, M.; Stahl, K.; Schluckebier, G. & Al-Karadaghi, S. Characterization of insulin microcrystals using powder diffraction and multivariate data analysis. *J. Appl. Cryst*. **2006**, *39*, 391–400.
51. Norrman, M. & Schluckebier, G. Crystallographic characterization of two novel crystal forms of human insulin induced by chaotropic agents and a shift in pH. *BMC Struct. Biol*. **2007**, *7*, 83.
52. Valmas, A.; Magiouf, K.; Fili, S.; Norrman, M.; Schluckebier, G.; Beckers, D.; Degen, T.; Wright, J.; Fitch, A.; Gozzo, F.; Giannopoulou, A.E.; Karavassili, F. & Margiolaki, I. Novel crystalline phase and first-order phase transitions of human insulin complexed with two distinct phenol derivatives. *Acta Cryst. D* **2015**, *71*, 819–828.

53. Fili, S.; Valmas, A.; Norrman, M.; Schluckebier, G.; Beckers, D.; Degen, T.; Wright, J.; Fitch, A.; Gozzo, F., Giannopoulou, A.E., Karavassili, F. & Margiolaki, I. Human insulin polymorphism upon ligand binding and pH variation: the case of 4-ethylresorcinol. *IUCrJ* **2015**, *2*, 534–544.
54. Pawley, G.S. Unit-cell refinement from powder diffraction scans. *J. Appl. Cryst.* **1981**, *14*, 357–361.
55. Smith, G.D.; Pangborn, W.A. & Blessing, R.H. Phase changes in $T_3R_3^f$ human insulin: temperature or pressure induced? *Acta Cryst*. D **2001**, *57*, 1091–1100.
56. Von Dreele, R.B.; Stephens, P.W.; Smith, G.D.; Blessing R.H. The first protein crystal structure determined from high-resolution X-ray powder diffraction data: A variant of T_3R_3 human insulin-zinc complex produced by grinding. *Acta Cryst. D* **2000**, *56*, 1549–1553.
57. Von Dreele, R.B. Binding of *N*-acetylglucosamine to chicken egg lysozyme: a powder diffraction study. *Acta Cryst*. D **2001**, *57*, 1836–1842.
58. Von Dreele, R.B. Binding of *N*-acetylglucosamine oligosaccharides to hen egg-white lysozyme: a powder diffraction study. *Acta Cryst. D* **2005**, *61*, 22–32.
59. Basso, S.; Fitch, A.N.; Fox, G.C.; Margiolaki, I. & Wright, J.P. High-throughput phase-diagram mapping *via* powder diffraction: a case study of HEWL versus pH. *Acta Cryst*. D **2005,** *61*, 1612–1625.
60. Krayenbühl, C. & Rosenberg, T. Crystalline protamine insulin. *Rep. Steno. Mem. Hosp. Nord. Insulinlab.* **1946**, 1, 60–73.
61. Balschmidt, P.; Hansen, F.B.; Dodson, E.J.; Dodson, G.G. & Korber, F. Structure of porcine insulin cocrystallized with clupeine Z. *Acta Cryst. B* **1991**, *47*, 975–986.
62. Cerenius, Y.; Stahl, K.; Svensson, L.A.; Ursby, T.; Oskarsson, A.; Albertsson, J. & Liljas, A. The crystallography beamline I711 at MAX II. *J. Synchrotron Rad*. **2000**, *7*, 203–208.
63. Mammen, C.B.; Ursby, T.; Cerenius, Y.; Thunnissen, M.; Als-Nielsen, J.; Larsen, S. & Liljas, A. Design of a 5-station macromolecular crystallography beamline at MAX- Lab. *Acta Phys. Pol. A*, **2002**, *101*, 595–602.
64. Wold S.; Esbensen K.; Geladi P. Principal component analysis. *Chemom. Intell. Lab.Syst.* **1987**, 2, 37-52., https://doi.org/10.1016/0169-7439(87)80084-9.
65. Smith, G.D.; Pangborn, W.A. & Blessing, R.H. The structure of T_6 human insulin at 1.0 Å resolution. *Acta. Cryst. D* **2003**, *59*, 474–482.
66. Margiolaki, I.; Giannopoulou, A.E.; Wright, J.P.; Knight, L.; Norrman, M.; Schluckebier, G.; Fitch, A.N.; Von Dreele, R.B. High-resolution powder X-ray data reveal the T6 hexameric form of bovine insulin. *Acta Cryst.* D **2013**, *69*, 978–990.
67. Larson, A.C.; & Von Dreele, R.B. General Structure Analysis System (GSAS). *Los Alamos National Laboratory Report LAUR*, **2004**, 86-748.
68. Toby, B.H. EXPGUI, a graphical user interface for GSAS. *J. Appl. Cryst*. **2001**, *34*, 210–213.
69. Bhat, T.N. Calculation of an OMIT map. *J. Appl. Cryst*. **1988**, 21, 279–281.
70. Rowe, R.C.; Sheskey, P.J. & Owen, S.C. *Handbook of Pharmaceuticals Excipients*, 5th ed.; Pharmaceutical Press: London, UK, 2005.
71. Matthews, B.W. Solvent content of protein crystals. *J. Mol. Biol*. **1968**, *33*, 491–497.
72. Kantardjieff, K.A. & Rupp, B. Matthews coefficient probabilities: Improved estimates for unit cell contents of proteins, DNA, and protein–nucleic acid complex crystals. *Protein Sci*. **2003**, *12*, 1865–1871.
73. Karavassili, F., *et al*. Novel monoclinic crystalline form of human Insulin complexed with Meta-cresol. *Acta Cryst. D* (under preparation).
74. Degen, T.; Sadki, M.; Bron, E.; König, U. & Nénert, G. The HighScore suite. *Powder Diffr*. **2014**, *29*, S13–S18.
75. 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, 5837–5845.
76. 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* **2001**, *62*, 249–260.
77. Bloom, C.R.; Choi, W.E.; Brzovic, P.S.; Ha, J.J.; Huang, S.T.; Kaarsholm, N.C. & Dunn, M.F. Ligand binding to wild-type and E-B13Q mutant insulins: A three-state allosteric model system showing half-site reactivity. *J. Mol. Biol*. **1995**, *245*, 324–330.
78. Rabinow, B.E. Nanosuspensions in drug delivery. *Nat. Rev. Drug Discov*. **2004**, *3*, 785–796.
79. Blair, A.S.; Hajduch, E.; Litherland, G.J.; Hundal, H.S. Regulation of glucose transport and glycogen synthesis in L6 muscle cells during oxidative stress. Evidence for cross-talk between the insulin and SAPK2/p38 mitogen-activated protein kinase signaling pathways. *J. Biol. Chem*. **1999**, *274*, 36293-36299.

80. Ceriello, A., & Motz, E. Is oxidative stress the pathogenic mechanism underlying insulin resistance, diabetes, and cardiovascular disease? The common soil hypothesis revisited. *Arterioscler. Thromb. Vasc. Biol*. **2004**, *24*, 816–823.
81. Baynes, J.W., Thorpe, S.R. Role of oxidative stress in diabetic complications: A new perspective on an old paradigm. *Diabetes* **1999**, 48, 1–9.
82. Johansen, J.S.; Harris, A.K.; Rychly, D.J.; Ergul, A. Oxidative stress and the use of antioxidants in diabetes: linking basic science to clinical practice. *Cardiovasc Diabetol.* **2005**, 4:5, doi: 10.1186/1475-2840-4-5.
83. Cathcart R.F. 3rd, Vitamin C: the nontoxic, nonrate-limited, antioxidant free radical scavenger, *Med. Hypotheses* **1985**, 18, 1, 1861–77.
84. Kodama, M.; Kodama, T.; Murakami, M.; Kodama, M. Diabetes mellitus is controlled by vitamin C treatment. *In Vivo* ***1993***, 7(6A), 535-542.
85. Dakhale, G.N.; Chaudhari, H.V.; Shrivastava, M. Supplementation of vitamin C reduces blood glucose and improves glycosylated hemoglobin in type 2 diabetes mellitus: A randomized, double-blind study. *Advances in Pharmacological Sciences* **2011**: article ID 195271.
86. Valmas, A.; Fili, S.; Kontou, P.; Norrman, M.; Schluckebier, G.; Beckers, D.; Degen, T.; Wright, J.P.; Fitch, A.N.; Karavassili, F. & Margiolaki. Conformational responses to pH of human insulin (under preparation).
87. Karavassili, F.; Valmas, A; Fili, S.; Kontou, P.; Spiliopoulou, M.; Bowler, M.; Beckers, D.; Degen, T.; Nenert, G; Fitch, A.N.; Wright, J.P.; Norrman, M.; Schluckebier G. and Margiolaki I. Microcrystalline cubic human insulin at 2.7Å resolution. *Acta Cryst. D* (under preparation).
88. Badger, J.; Harris, M.R.; Reynolds, C.D. & Evans, A. C. Structure of pig insulin dimer in the cubic crystal. *Acta Cryst*. B **1991**, 47, 127–135.
89. Yu, B., & Caspar, D.L.D. Structure of Cubic Insulin Crystals in Glucose Solutions. *Biophys. J.* **1998**, 74, 616 –622.
90. Scott, D.A. Crystalline insulin. *Biochem. J.* **1934**, 28, 4, 1592–1602.
91. Edsall, J.T.; Felsenfeld, G.L.D.; Dewitt, S.; Goodman, D.S. & Gurd, F.R.N. The Association of Imidazole with the Ions of Zinc and Cupric Copper[1a,b,c]. *J. Amer. Chem. Soc.* **1954**, 76, 3054-3061.
92. Rotenstein, L.S.; Ran, N.; Shivers, J.P.; Yarchoan, M.; Close, K.L. Opportunities and challenges for biosimilars: What's on the horizon in the global insulin market? *Clin. Diabetes* **2012**, 30, 4, 138-150.
93. Bernstein, J. *Polymorphism in Molecular Crystals*; Clarendon Press: Oxford, UK, 2002.