Acta Pharmaceutica Sinica B 2025;15(11):5880−5890

Chinese Pharmaceutical Association
Institute of Materia Medica, Chinese Academy of Medical Sciences

Acta Pharmaceutica Sinica B

www.elsevier.com/locate/apsb
www.sciencedirect.com

ORIGINAL ARTICLE

Deciphering the significant impact of natural glycosylation on human insulin

Yaohao Li[a], Wenqiang Liu[a], Dan Liu[a], Ruihan Wang[a,b], Yajing Zhang[a], Xin Li[a], Jinyuan Gong[a], Shiying Shang[c], Zhongping Tan[a,*]

[a]*State Key Laboratory of Bioactive Substance and Function of Natural Medicines, Institute of Materia Medica, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100050, China*
[b]*Chemical Engineering College, Hebei Normal University of Science and Technology, Qinhuangdao 066600, China*
[c]*Center of Pharmaceutical Technology, School of Pharmaceutical Sciences, Tsinghua University, Beijing 100084, China*

Received 2 January 2025; received in revised form 26 March 2025; accepted 29 May 2025

KEY WORDS

Insulin;
Insulin synthesis;
Glycosylation;
Rapid-acting;
Solubility

Abstract In the century-long evolution of insulin pharmaceuticals, each transformative advancement in this drug class has been closely tied to the ability to obtain new insulin isoforms for research. Despite this, the recently discovered naturally occurring isoforms of glycosylated human insulin have remained largely unattainable for proper characterization. Herein, we demonstrate for the first time that total chemical synthesis can be used to generate all isoforms. This achievement required maintaining the correct positions of the interchain disulfide bonds while effectively removing protecting groups on complex glycans. Notably, the availability of seven glycoforms reveals the important effects of natural sialylated glycans in suppressing insulin self-association and enhancing its solubility, surpassing the performance of currently employed rapid-acting insulin drugs. This work not only offers a readily adaptable platform for exploring natural *O*-glycosylation in other therapeutic proteins and peptides but also lays the groundwork for further research into harnessing natural glycosylation for therapeutic applications.

*Corresponding author.
E-mail address: zhongping.tan@imm.ac.cn (Zhongping Tan).

Peer review under the responsibility of Chinese Pharmaceutical Association and Institute of Materia Medica, Chinese Academy of Medical Sciences.

https://doi.org/10.1016/j.apsb.2025.08.005

1. Introduction

For over a century, insulin research has played a pioneering role in developing protein drugs and delivery systems, refining techniques for protein analysis and synthesis, innovating technology for protein expression and modification, and advancing our understanding of protein structure and biology[1,2]. In particular, insulin research has made transformative contributions to diabetes treatment, saving countless lives and empowering millions to effectively manage the disease[3,4]. These remarkable achievements are intricately connected to the relentless pursuit of insulin isoforms with enhanced performance. The journey began with the groundbreaking discovery and clinical application of animal insulins in the 1920s by Banting et al.[5], revolutionizing diabetes treatment from a fatal disease to a manageable chronic condition (Fig. 1A). In the late 1970s, the advent of recombinant human insulin by Genentech greatly improved therapy accessibility and reduced immunogenicity[6] (Fig. 1B). Since the late 20th century, mutated and synthetically modified insulin analogs and their premixed formulations have enabled a closer approximation of normal insulin profiles, enhancing treatment efficacy[7,8] (Fig. 1C). These evolving advancements underscore the importance of continued creation of new insulin isoforms for further enhancing diabetes treatment. Consistent with the transition from animal insulins to human insulin, shifting from isoforms with unnatural elements (*e.g.*, mutations or modifications) to naturally occurring modifications might represent a logical direction for exploration (Fig. 1D).

In 2017, Li et al.[9] serendipitously discovered naturally occurring oxygen atom-linked glycosylation (*O*-linked glycosylation) at Thr27 of the B chain of human insulin (ThrB27) *via* mass spectrometry, identifying the glycan structures as HexNAc(1), HexNAc(1)Hex(1), HexNAc(2)Hex(1), HexNAc(1)Hex(1)NeuAc(1), and HexNAc(1)Hex(1)NeuAc(2). This finding was supported by Schjoldager et al.[10] in 2020, suggesting the glycans are mucin-type. Integrating literature data, it is inferred that the naturally occurring insulin glycoforms correspond to structures **Ins1**−**3** and **Ins5**−**7**[11,12] (Fig. 1D). Glycosylation of other therapeutic proteins has been found to modulate properties such as aggregation and self-association (*e.g.*, interferon-β)[13], solubility (*e.g.*, α-galactosidase A)[14], and stability (*e.g.*, erythropoietin)[15]. These effects have spurred interest in investigating the impact of chemical glycosylation of insulin, yielding encouraging results that underscore the importance of exploring the recently discovered naturally occurring *O*-glycosylation[16-18]. However, a significant barrier to such research is the lack of the glycoforms shown in Fig. 1D. Considering that these glycoforms originate from human islets, synthesis appears to be a more feasible approach than isolation for obtaining them at present. Both enzymatic and chemical synthesis methods show promise in generating insulin glycoforms. However, while enzymatic methods enable the rapid generation of complex glycoforms, they are currently limited by

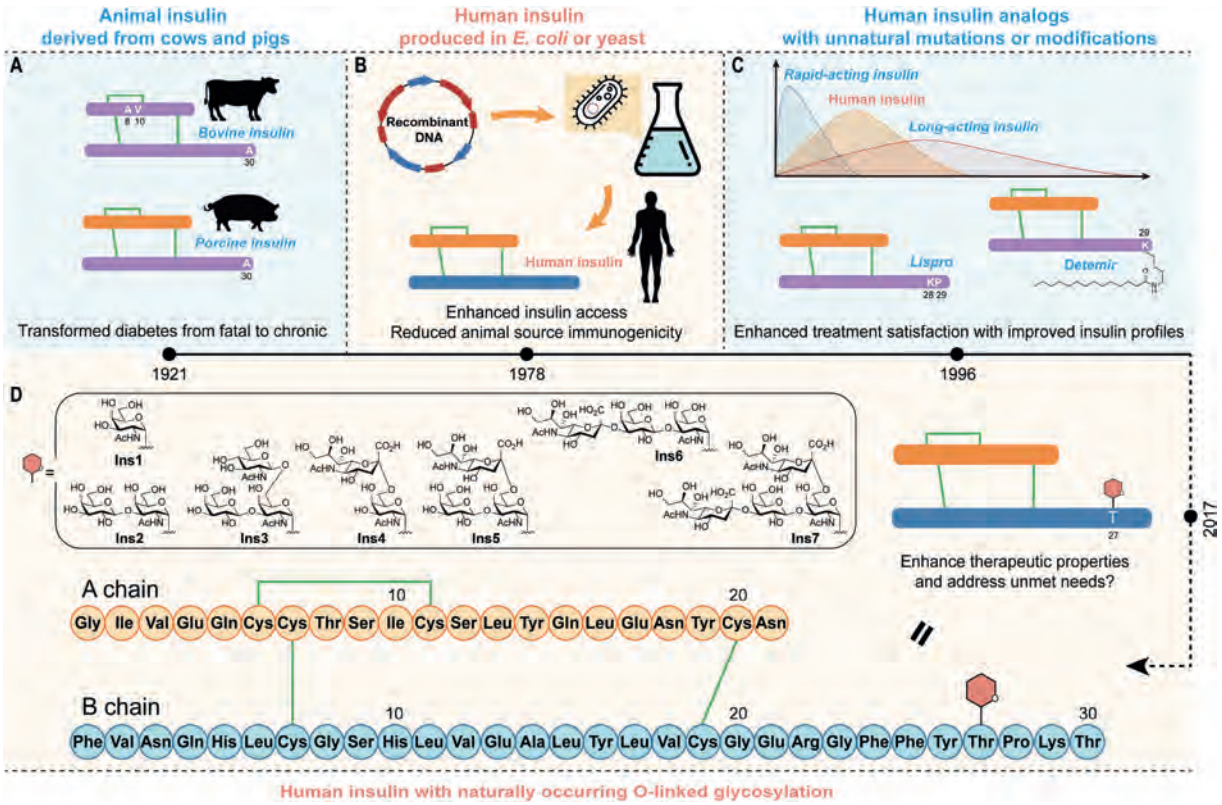


Figure 1 The natural glycosylation of human insulin poses a research challenge while also indicating a new direction for insulin pharmaceuticals, aligning with the long-standing pursuit of more effective diabetes management. Breakthroughs in insulin development, including its isolation from animals by Banting et al.[5] (A), biosynthetic production (B), and the creation of analogs like Lispro and Detemir (C), have significantly enhanced treatment options. Exploration of the natural glycosylation of human insulin (D) holds promise for further enhancing its therapeutic properties and addressing previously unmet needs. Disulfide bonds are depicted in green, while chains with sequences differing from human insulin are shown in purple.

scalability, producing insufficient material for detailed studies. In contrast, chemical synthesis allows for scalable glycoform preparation to investigate glycosylation effects, though it remains less efficient and requires significant optimization for generating complex glycoforms[19].

Historically, there have been numerous studies of chemical insulin synthesis[20]. However, despite significant advancements, constructing complex molecules like **Ins5−7** remains challenging[21]. Retrosynthetic strategies for assembling these glycoforms remain uncertain due to the unresolved key issue of effectively removing protecting groups on the sialylated *O*-glycans while ensuring the proper formation and preservation of delicate disulfide bonds (Fig. 2). The goal of this study was to develop a generally applicable strategy for the synthesis of glycosylated insulin, thereby enabling the investigation of its naturally occurring *O*-glycosylation. The artificial glycoform, **Ins4**, was also included in the study to complete the series alongside naturally occurring glycoforms (Fig. 1D). The systematic variation in glycan structures among these glycoforms has allowed us to pinpoint key glycan features that are essential for modulating insulin's therapeutic properties.

2. Results and discussion

2.1. *Total synthesis of insulin glycoforms*

Drawing inspiration from the logic behind the recently developed synthesis of glycosylated insulins[16,18], we initially explored a final-step glycan deprotection strategy for insulin glycoforms (Fig. 3, strategy I), anticipating that appropriate conditions would completely remove all glycan protecting groups while preserving the integrity of the disulfide bonds. The application of this strategy was exemplified through the total synthesis of **Ins1**, which began with the preparation of the isoacyl A chain, **InsA**, utilizing a Rink Amide AM resin and key building blocks: an *O*-acyl isodipeptide *Boc*-Ser[Fmoc-Thr(tBu)]-OH, differently protected Cys residues Fmoc-Cys(Trt)-OH, Fmoc-Cys(Mmt)-OH, Fmoc-Cys(Acm)-OH, and Fmoc-Cys(STmp)-OH, with the aid of microwave-assisted solid-phase peptide synthesis (SPPS) (Fig. 2 and Supporting Information Fig. S1). The previously validated isoacyl structure, used as a surrogate for Thr8−Ser9[22], along with the newly adopted Fmoc-Cys(STmp)-OH, which has a more easily removable side chain[23], and the more efficient microwave-assisted SPPS, addresses the low yield and handling issues encountered in prior syntheses of the A chain. After SPPS, the STmp group on CysA6 is selectively removed, and the resulting thiol group is activated with 2,2′-dithiobis(5-nitropyridine) (DTNP) on resin. Notably, during the side chain deprotection and peptide cleavage from the resin, the desired intramolecular disulfide bond CysA6−CysA11 is established automatically and selectively, thereby obviating the need for the separate formation of this disulfide bond. A single high-performance liquid chromatography (HPLC) purification afforded **InsA** in 21% yield (Fig. S1).

Ins1B was synthesized using Fmoc-Thr(tBu)-2-ClTrt resin, a cost-effective alternative to the previously used pre-loaded HMPB ChemMatrix® resin[22], reducing expenses from approximately $100 to $10 per gram without compromising efficiency (Fig. 2). The Cys building blocks employed in the synthesis were Fmoc-Cys(Trt)-OH and Fmoc-Cys(Acm)-OH. The Fmoc-protected glycosylated Thr, **FPGA1**, was utilized during SPPS to site-selectively introduce glycan onto ThrB27. Following SPPS, peptide deprotection and cleavage, HPLC purification afforded **Ins1B** in 28% yield (Supporting Information Fig. S2). The pivotal chain-combination step was executed by mixing **InsA** and **Ins1B** in a pH 8.0 buffer, which

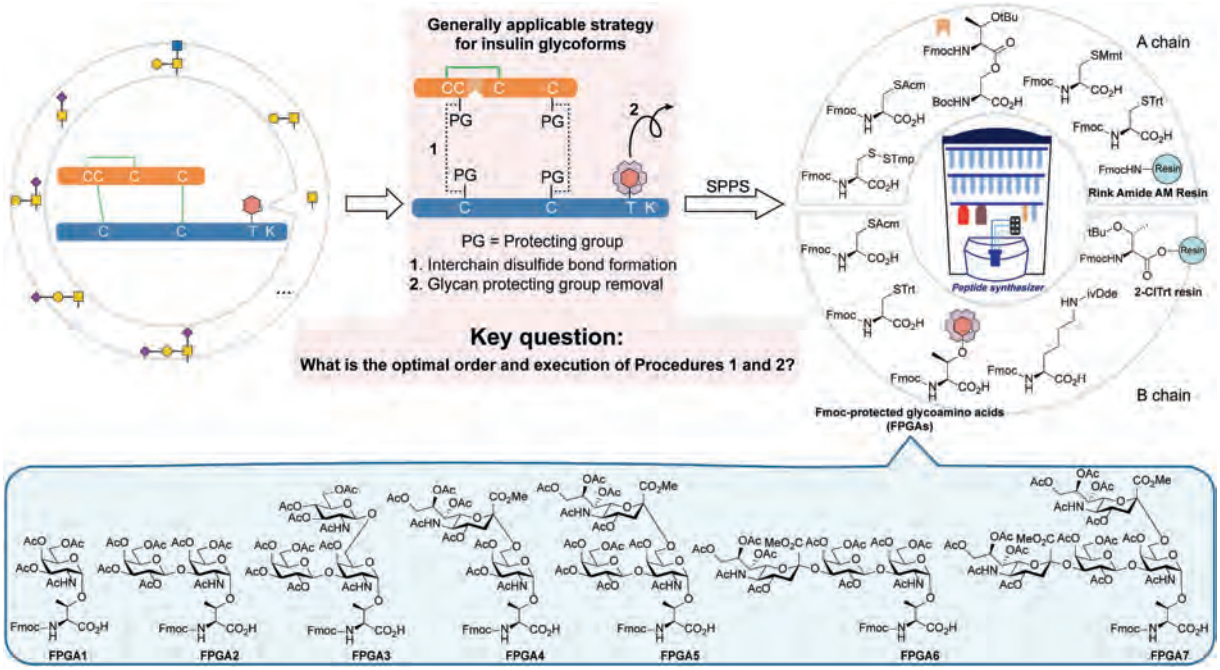
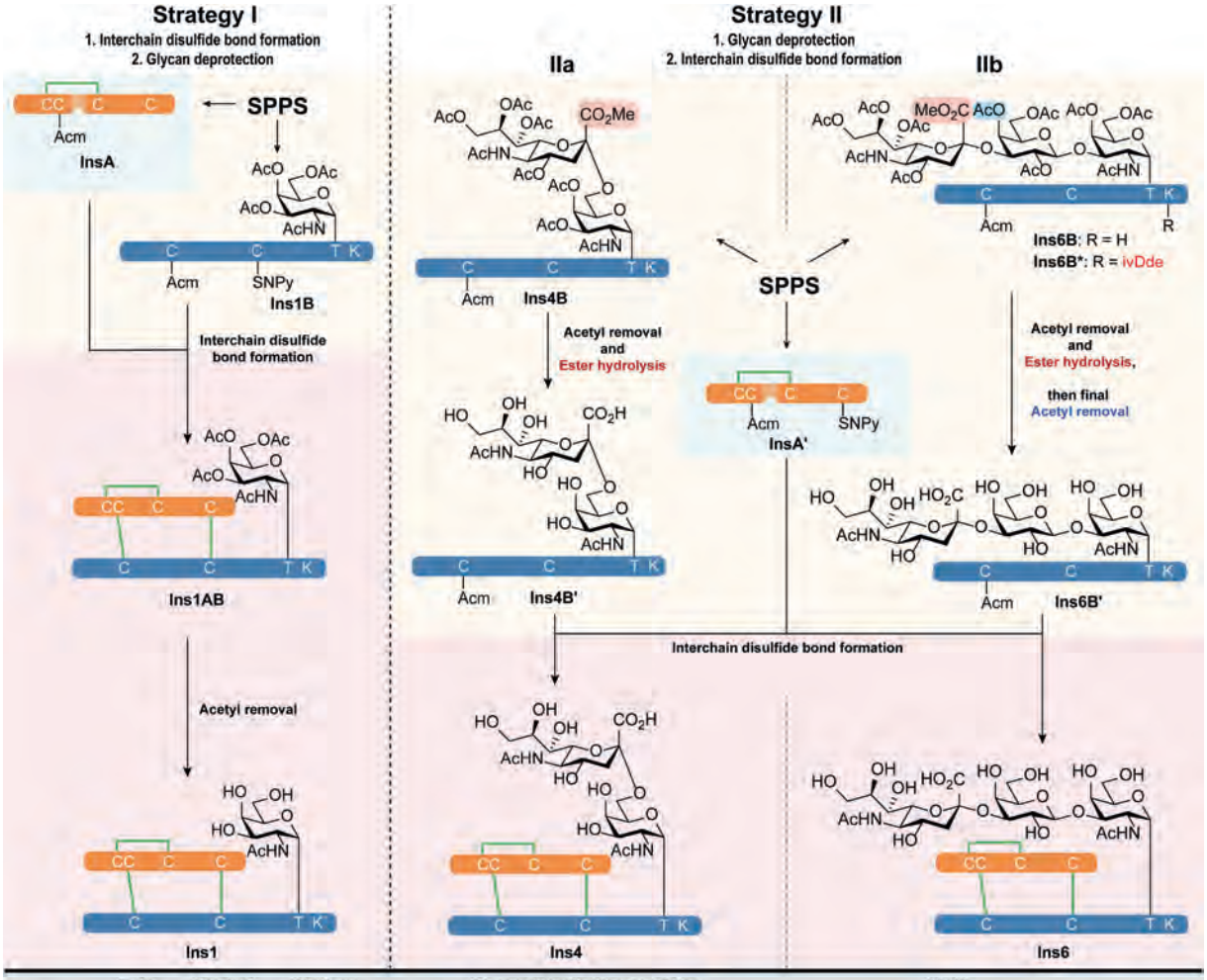


Figure 2 Retrosynthetic analysis for *O*-glycosylated human insulin. Developing a generally applicable synthetic strategy to obtain insulin *O*-glycoforms required addressing the major challenge of correctly forming disulfide bonds while successfully removing protecting groups necessary for the SPPS of the B chain. Tmp, 2,4,6-trimethoxyphenyl; Acm, acetamidomethyl; Mmt, 4-methoxytrityl; Trt, trityl; 2-ClTrt, 2-chlorotrityl; ivDde, 1-(4,4-dimethyl-2,6-dioxocyclohexylidene)-3-methylbutyl.



Entry	Product	Strategy	Glycan deprotection condition	Yield
1	**Ins1**	I	5% aq. NH_2NH_2, 0.5 h	SPPS: 28%; Com: 39%; Dep: 80%.
2	**Ins2**	I	5% aq. NH_2NH_2, 0.5 h	SPPS: 20%; Com: 36%; Dep: 88%.
3	**Ins3**	I	5% aq. NH_2NH_2, 0.5 h	SPPS: 26%; Com: 35%; Dep: 82%.
4	**Ins4**	IIa	30 mmol/L aq. NaOH, 4 °C, 6 h	SPPS: 25%; Dep: 37%; Com: 47%.
5	**Ins5**	IIa	30 mmol/L aq. NaOH, 4 °C, 6 h	SPPS: 24%; Dep: 33%; Com: 35%.
6[a]	**Ins6**	IIb	1) 40 mmol/L aq. NaOH, 1 h; 2) 15% aq. NH_2NH_2, 3h	SPPS: 25%; Dep: 43%; Com: 31%.
7[a]	**Ins7**	IIb	1) 40 mmol/L aq. NaOH, 1 h; 2) 15% aq. NH_2NH_2, 3h	SPPS: 19%; Dep: 32%; Com: 37%.
8[b]	**Ins6**	IIb	1) 40 mmol/L aq. NaOH, 1 h; 2) 15% aq. NH_2NH_2, 3h	SPPS: 24%; Dep: 58%; Com: 31%.
9[b]	**Ins7**	IIb	1) 40 mmol/L aq. NaOH, 1 h; 2) 15% aq. NH_2NH_2, 3h	SPPS: 18%; Dep: 50%; Com: 37%.

More general

Less complex

Figure 3 Categorization and development of strategies for the total synthesis of *O*-glycosylated human insulin. Strategy I, being less complex, is optimal for glycoforms without sialic acids, such as **Ins1**−**3**. Strategy IIa is more appropriate for glycoforms containing α2,6-linked sialic acid, like **Ins4** and **5**. The development of the more complex but broadly applicable Strategy IIb marked a breakthrough, allowing for the efficient removal of challenging acetyl groups while minimizing side product formation due to acetyl migration. [a]Insulin glycoforms synthesized using **Ins6B** and **Ins7B**. [b]Insulin glycoforms synthesized using **Ins6B*** and **Ins7B***. Abbreviations: NPy, 5-nitropyridin-2-yl-.

facilitated the formation of the CysA20−CysB19 disulfide bond. This process triggered the intramolecular acyl migration reaction within the isoacyl structure, thereby restoring the native peptide bond. Subsequently, the second interchain disulfide bond, CysA7−CysB7, was formed through the *in situ* addition of I_2, affording **Ins1AB** in 39% yield (Fig. 3 and Supporting Information Fig. S3). The final product, **Ins1**, was obtained by removing the acetyl (Ac) groups with NH_2NH_2, which proved to be more efficient in cleaving protecting groups and preserving disulfide linkages compared to other commonly used reagents like NaOH, NaOMe, and $NH_3 \cdot H_2O$ (Supporting Information Figs. S4 and S5). The yield for this step was 80%, and the purity of **Ins1** was rigorously validated using ultra-performance liquid chromatography (UPLC), which displayed a sharp, single peak (Fig. S5). High-resolution mass spectrometry (HRMS) confirmed the identity of **Ins1**, showing the expected *m*/*z* signals (Fig. S5). The correct disulfide connectivity was verified through Glu-C digestion analysis. Furthermore, the proper folding of **Ins1** was confirmed by circular dichroism (CD), which exhibited spectra closely matching that of unglycosylated insulin. As expected, **Ins2** and **Ins3** were also effectively obtained following strategy I with yields comparable to that of **Ins1** (Fig. S5).

We then turned our attention to preparing **Ins4**, a synthesis that could be complicated by the presence of the α2,6-linked sialic acid. Initial tests confirmed that the NH_2NH_2 solution used in strategy I reacted with the methyl (Me) ester, leading to the formation of an undesired hydrazide (Supporting Information Fig. S6A). Consequently, $NH_3 \cdot H_2O$ was chosen for deprotection due to its proven effectiveness in both Ac group removal and methyl ester hydrolysis[18]. However, this approach resulted in an amide by-product with a −1 Da mass shift (Fig. S6B). Since the use of NaOH or NaOMe solutions or their combinations led to undesired disulfide bond rearrangements (Fig. S4), a new strategy, strategy IIa, was developed to overcome this synthetic challenge. This involved preparing a new A chain variant, **InsA′**, with a SNPy-activated CysA20, and a new B chain variant, **Ins4B′**, with an inactivated CysB19 and a fully unprotected glycan moiety (Fig. 3). **InsA′** was synthesized similarly to **InsA** (Fig. S1). **Ins4B′** was obtained with a 37% yield by treating the glycopeptide **Ins4B** with an optimized 30 mmol/L NaOH solution at 4 °C (Supporting Information Figs. S7 and S8). With the newly prepared **InsA′** and **Ins4B′**, selective and sequential formation of the two interchain disulfide bonds gave **Ins4** in 47% yield (Supporting Information Fig. S9). Similarly, **Ins5** was obtained using strategy IIa with a 35% yield (Fig. S9).

Unexpectedly, attempts to deprotect **Ins6B** using NaOH solution yielded an undesired by-product with incompletely removed Ac groups, resulting in an inseparable mixture with **Ins6B′** (Supporting Information Fig. S10A). Mass spectrometry fragmentation analysis indicated that the most challenging Ac group to remove from the α2,3-linked sialic acid-containing trisaccharide was likely located at the C4 position of the galactose (Gal) residue[24]. Similar outcomes were observed when using NaOMe (Fig. S10B). To circumvent this challenge, we opted for a trisaccharide with a cyclic structure between sialic acid and Gal, thereby eliminating the problematic Ac group[25]. Unfortunately, tests revealed that this trisaccharide could not tolerate the presence of TFA (Fig. S10C). Given the extraordinary difficulty of removing all protecting groups in a single step, we pivoted to a two-step deprotection strategy, strategy IIb (Fig. 3). After extensive optimization, we identified that a combination of 40 mmol/L NaOH and 15% NH_2NH_2 aqueous solutions was effective in converting **Ins6B** to **Ins6B′** (Fig. 3 and Supporting Information Figs. S11−S13). The initial NaOH hydrolysis completely removed the methyl ester, and after a preliminary HPLC separation, the subsequent NH_2NH_2 treatment effectively removed all remaining Ac groups, yielding **Ins6B′** in 43% yield. Despite this success, the deprotection process also generated a by-product with an unremovable Ac group. Mass spectrometry fragmentation analysis suggested that this by-product might result from Ac group migration to LysB29 during NaOH-mediated hydrolysis (Figs. S11 and S12). To verify this insight, we replaced the Boc protecting group on LysB29 with the NaOH-stable and hydrazine-labile ivDde. This strategic modification effectively circumvented the formation of the by-product, resulting in a 58% yield of **Ins6B′** from **Ins6B*** (Fig. 3 and Fig. S13). Similarly, **Ins7B′** was synthesized from **Ins7B** or **Ins7B*** with yields of 32% and 50%, respectively. **Ins6** and **Ins7** were readily synthesized by combining **Ins6B′** and **Ins7B′** with **InsA′** using the developed procedure, achieving yields of 31% and 37%, respectively (Supporting Information Fig. S14).

In this study, insulin glycoforms were synthesized at a 10 mg scale due to the high cost of FPGAs. This quantity was sufficient for comprehensive characterization. The optimized synthetic strategy is expected to be scalable for industrial applications. With the commercial availability of FPGAs at a reasonable cost, the A and B chains could be synthesized and assembled on a larger scale, similar to the industrial production of Fuzeon, a 36-amino acid anti-HIV peptide drug manufactured *via* SPPS[26] (Supporting Information Figs. S15−S18).

2.2. *Properties of insulin glycoforms*

With these glycoforms in hand, we initially evaluated the impact of glycosylation on the self-association of insulin, a critical factor in regulating its bioavailability and the timing and intensity of its hypoglycemic effect[27,28]. Unglycosylated human insulin is known to have a strong propensity for self-association, forming stable hexamers in the presence of Zn^{2+} ions, resulting in slow absorption from subcutaneous tissue into the bloodstream[28]. In contrast, its mutant analogs, including Aspart, Lispro, and Glulisine, exhibit a weaker tendency for self-association, leading to a faster absorption rate and onset of action, and are currently utilized as rapid-acting therapies[8,29] (Fig. 1).

Analytical ultracentrifugation (AUC), a classical technique for studying protein aggregation in solution, was employed to investigate how insulin glycosylation influences its self-association behavior[30]. For comparison, commercially purchased and purified unglycosylated human insulin, Aspart, Lispro, and Glulisine were included in the study. The self-association of insulins was initially assessed in phosphate-buffered saline (PBS) without Zn^{2+}, at a concentration of 60 μmol/L, which is 1/10 of the 100 units/mL commonly used for diabetes treatment. Under this condition, the unglycosylated insulin demonstrates a stronger tendency for self-association compared to both the mutant and glycosylated insulins, with respective monomer, dimer, and hexamer ratios of 45.3%, 42.2%, and 12.5% (Fig. 4A and Supporting Information Fig. S19). Glycosylation leads to an increase in monomer ratios and a decrease in dimer ratios, a trend closely correlated with the number of glycan units. **Ins3** and **Ins5−7**, which contain tri- or tetrasaccharides, exhibit monomer ratios (exceeding 80%) comparable to those of Aspart and Lispro. Interestingly, Glulisine demonstrates relatively low monomer and high hexamer ratios.

The impact of glycosylation on insulin self-association was next evaluated in a typical buffer used for rapid-acting insulin formulations[31], containing Zn^{2+} at 60 μmol/L. As shown in Fig. 4A and Supporting Information Fig. S20, the advantages of glycosylation begin to emerge in the presence of Zn^{2+}. Unglycosylated human insulin, Aspart, Lispro, and Glulisine now show a tendency for self-association comparable to glycoforms **Ins1−4**, mainly forming hexamers and exhibiting monomer ratios lower than 45%. However, **Ins5−7**, which bear sialylated tri- or tetrasaccharides, maintain high monomer ratios of about 80%. These findings indicate that, under these experimental conditions, glycosylation with relatively large *O*-linked sialylated glycans significantly reduces insulin self-association, surpassing the effects of previously adopted mutations.

To challenge this observation further, we examined insulin self-association under more stringent conditions by increasing the concentration to 600 μmol/L (100 units/mL)[31]. Under these conditions, unglycosylated insulin forms, including Aspart, Lispro, and Glulisine, exhibited negligible monomer ratios (<3%), with Aspart and Lispro primarily forming hexamers (>90%), while Glulisine mainly existed as di-hexamers (52.8%) (Fig. 4A and Supporting Information Fig. S21). Glycosylation conferred resistance to multimer formation at the higher insulin concentration.

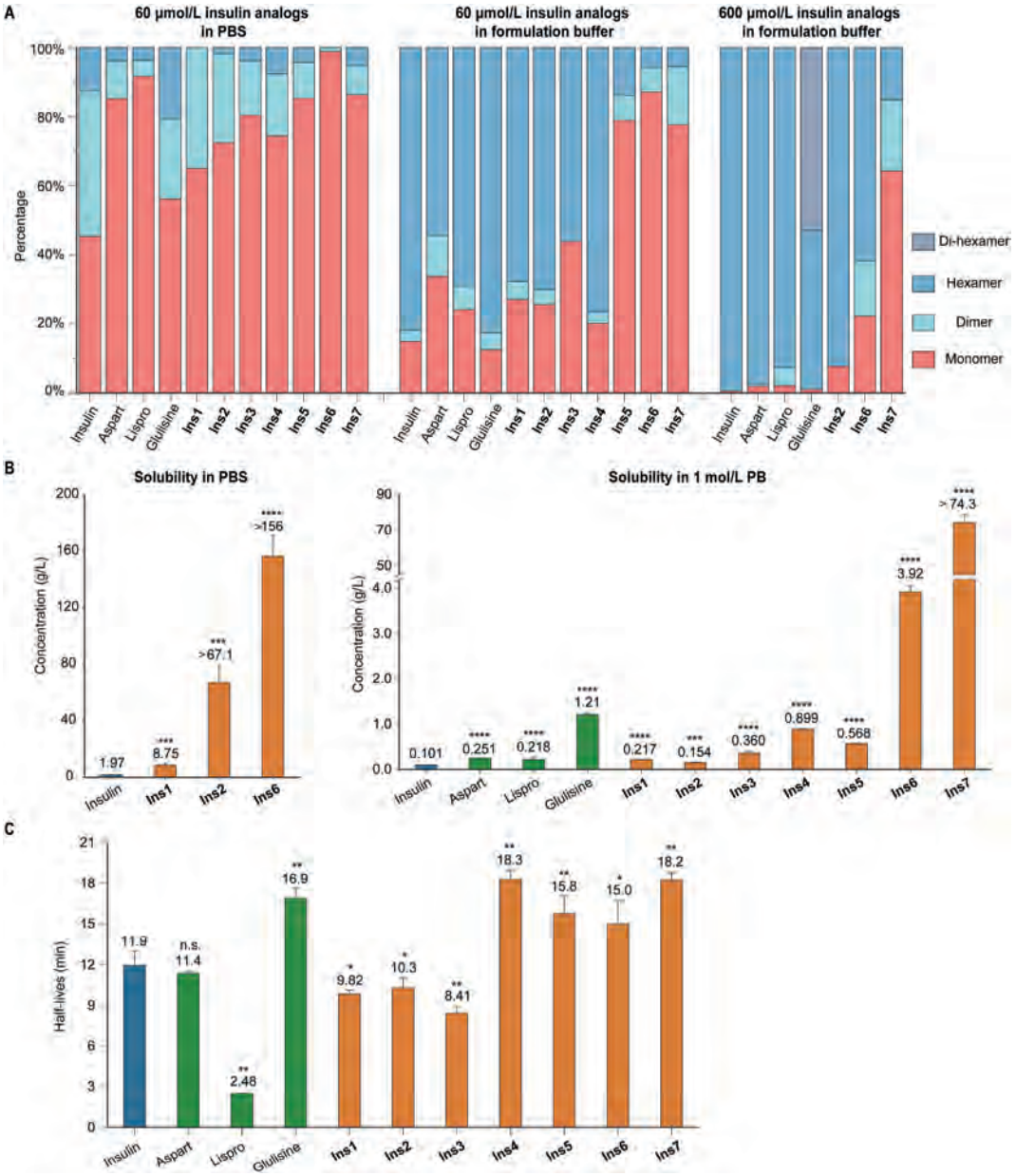


Figure 4 The impact of glycosylation on insulin properties varies from negligible to significant depending on sialylation levels. (A) Self-association tendencies of insulin in PBS (60 μmol/L) and in insulin formulation (16 g/L glycerol, 1.50 g/L phenol, 1.72 g/L meta-cresol, 19.6 mg/L Zn^{2+}, 1.25 g/L $Na_2HPO_4 \cdot 2H_2O$, and 0.58 g/L NaCl, pH 7.4) at concentrations of 60 and 600 μmol/L. Monomers, dimers, hexamers, and di-hexamers are depicted in red, cyan, blue, and slate blue, respectively, with the percentages of each species indicated by the sector height. (B) Solubility of insulin variants in PBS and in 1 mol/L PB (pH 7.4). (C) Half-lives of insulin variants determined by chymotrypsin digestion (0.5 g/L) in a buffer containing 100 mmol/L Tris and 1.0 mmol/L $CaCl_2$ (pH 8.0). Data in (B) and (C) are presented as the mean ± SD ($n = 3$). P values were calculated using a two-tailed Student's t-test. n.s., $P > 0.05$; *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$; ****$P < 0.0001$, compared with unglycosylated human insulin.

The unsialylated disaccharide on **Ins2** noticeably increased the monomer ratio to about 8%. The sialylated trisaccharide on **Ins6** further enhanced the monomer ratio to 22.3%. Impressively, the disialylated tetrasaccharide on **Ins7** increased the monomer ratio to over 60%, with the hexamer ratio reduced to only 15.2%. Taken together, these findings confirm the extraordinary capability of sialylated *O*-glycans to prevent insulin self-association. The exceptionally high monomeric ratio of glycosylated insulins, particularly **Ins6** and **Ins7**, suggests they may rapidly dissociate into monomers upon injection or even bypass the rate-limiting dissociation step. These glycosylated insulins hold strong potential for developing enhanced rapid-acting or ultra-rapid-acting formulations, offering faster onset of action, greater dosing flexibility, and an improved option for patients struggling to achieve postprandial glycemic control with current bolus insulins[8,29].

Next, we explored whether glycosylation enhances insulin solubility[32], potentially enabling higher concentration formulations[33]. We first assessed insulin solubility by measuring their saturated

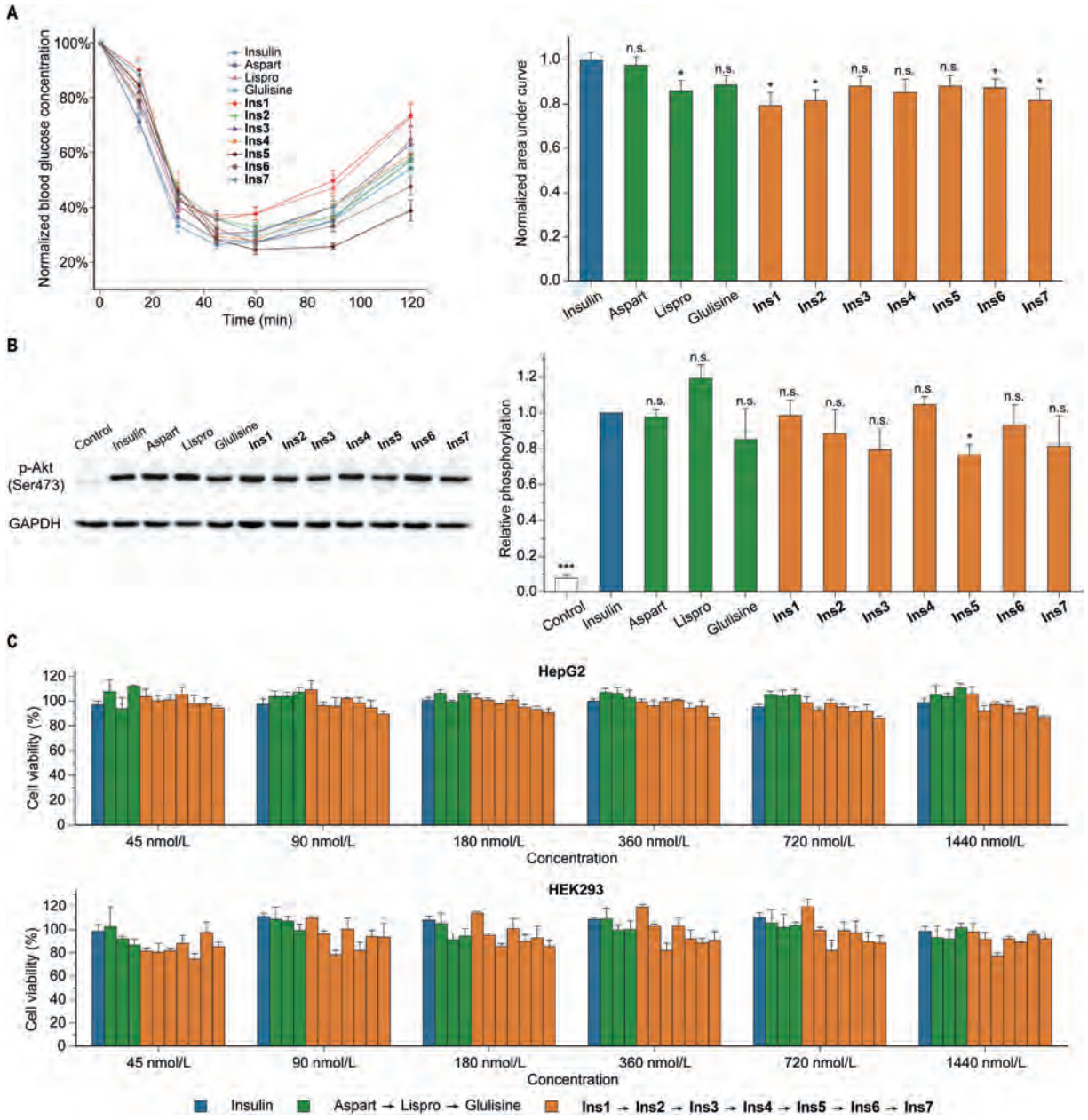


Figure 5 The impact of glycosylation on insulin functions is minimal. (A) Normalized blood glucose concentration over a 2-h period following subcutaneous administration of insulin variants (0.75 units/kg) in STZ-treated mice, along with the normalized area under the blood glucose curve for the first 1 h. Data are presented as mean ± SD ($n = 7$). (B) Immunoblot analysis and quantification of AKT phosphorylation at Ser 473 in HepG2 cells stimulated with 25 nmol/L insulin variants for 20 min. GAPDH was used as a loading control. Data are presented as mean ± SD ($n = 3$). (C) Cell viability of HepG2 and HEK293 cells exposed to varying concentrations of insulin variants. Data are presented as mean ± SD ($n = 3$). P values were calculated using a two-tailed Student's t-test. n.s., $P > 0.05$; *$P < 0.05$; ***$P < 0.001$, compared with unglycosylated human insulin.

concentrations in PBS[34] (Supporting Information Fig. S22 and Supporting Information Table S1). The results revealed that unglycosylated insulin had a solubility of approximately 1.97 g/L (57 units/mL; Fig. 4B). Glycosylation greatly increased insulin solubility. While **Ins 1** exhibited a modest four-fold increase (8.75 g/L, 244 units/mL), **Ins2** and **Ins6** showed dramatically higher solubilities of over 67 g/L and 156 g/L, respectively. Their exact solubility values were not determined due to the large sample mass required. To facilitate comparisons across all insulin isoforms, we intentionally reduced solubility using a 1 mol/L phosphate buffer (PB)[35], which lowered the solubility of unglycosylated insulin, **Ins1**, and **Ins2** to approximately 0.1−0.2 g/L, and **Ins6** to 3.92 g/L. In this buffer, Aspart and Lispro had solubilities of about 0.2 g/L, while Glulisine reached 1.21 g/L. Notably, sialylated large glycans greatly enhanced insulin solubility, with **Ins7** well exceeding 74 g/L. These results clearly demonstrate the ability of large sialylated *O*-linked glycans to increase insulin solubility. The high solubility of glycosylated insulins, particularly **Ins6** and **Ins7**,

allows for the development of insulin formulations at much higher concentrations than those achievable with currently marketed insulin drugs. This advancement is expected to increase insulin device capacity, enabling longer-lasting insulin pens, while also benefiting patients with severe insulin resistance who require high daily doses. By reducing injection volume and frequency, this improved solubility has the potential to enhance treatment compliance[33].

We also investigated the role of glycosylation in enhancing the proteolytic stability of insulin, which could potentially lead to a reduction in injection frequency[36,37]. Using chymotrypsin, we identified isoform-specific effects (Fig. 4C and Supporting Information Table S2)[38]. Aspart, Lispro, and **Ins1−3** exhibited lower half-lives than unglycosylated human insulin, with Lispro showing about a fivefold reduction and others only slightly lower. Conversely, Glulisine and insulins with sialylated glycans (**Ins4−7**) demonstrated increased stability. Among them, **Ins4** and **Ins7** exhibited the highest stability, with half-lives approximately 1.5 times that of unglycosylated human insulin. These findings suggest a modest positive impact of glycosylation on insulin proteolytic stability.

To evaluate long-term storage stability, we analyzed two representative glycosylated insulins, **Ins6** and **Ins7**, in PBS (60 μmol/L) and rapid-acting insulin formulation (600 μmol/L). Both glycosylated and unglycosylated insulins remained stable over 3 weeks at room temperature (Supporting Information Fig. S23, Supporting Information Tables S3 and S4).

2.3. *Functions of insulin glycoforms*

To determine whether glycosylation diminishes or alters the biological functions of insulin[39], we investigated the impact of *O*-glycosylation on glucose uptake using a streptozotocin (STZ) mouse model of type 1 diabetes[40]. Mice were administered insulin variants subcutaneously at a dose of 0.75 units/kg, and their blood glucose levels were monitored over a 2-h period. Encouragingly, our results showed that glycosylation had minimal impact on glucose uptake, preserving insulin's effectiveness in this essential function (Fig. 5A and Supporting Information Tables S5−S7).

We then explored the influence of *O*-glycosylation on AKT phosphorylation at the Ser473 residue, a key marker of PI3K/AKT signaling pathway activation, using HepG2 cells[41,42]. The cells were stimulated with 25 nmol/L of various insulin variants for 20 min, and AKT phosphorylation levels were assessed *via* immunoblot analysis. Our findings revealed that all glycosylated insulin variants activated AKT to a similar extent as unglycosylated insulin, indicating that glycosylation does not significantly alter this critical signaling pathway for insulin function (Fig. 5B, and Supporting Information Fig. S24 and Supporting Information Table S8). To further confirm this, we examined AKT pathway activation using two representative glycoforms, **Ins6** and **Ins7**, at varying concentrations (6.25, 12.5, 25, 50, and 100 nmol/L) and incubation times (2, 5, 10, 20, and 40 min). The results consistently showed that the phosphorylation level at Ser473 induced by glycosylated insulins was comparable to that of unglycosylated insulin (Supporting Information Fig. S25, Supporting Information Tables S9 and S10). Moreover, at a concentration of 25 nmol/L and an incubation time of 20 min, **Ins6** and **Ins7** also stimulated AKT phosphorylation at Thr308 to the same level as unglycosylated insulin (Supporting Information Fig. S26 and Supporting Information Table S11).

While insulin primarily activates the PI3K/AKT pathway, it also modulates additional signaling cascades, such as the CRK/TC10 pathway (regulating metabolic responses) and the Ras/MAPK pathway (mediating mitogenic effects)[43,44]. Although our data indicate that glycosylation minimally impacts PI3K/AKT signaling, we acknowledge the possibility of biased signaling, where glycosylation might selectively enhance or suppress specific pathways. Moreover, glycosylation may modulate interactions with receptors, such as the IGF-1 receptor, potentially contributing to off-target effects through differential receptor activation[45]. These important considerations will be systematically investigated in our ongoing research.

Finally, we evaluated the cytotoxic potential of *O*-glycosylated insulin by incubating HepG2 and HEK293 cells with insulin variants at concentrations ranging from 45 nmol/L to 1440 nmol/L[46]. Cell viability was assessed using the CCK-8 assay, comparing treated cells to a control culture medium. The data showed no significant cytotoxicity associated with glycosylated insulin variants at these concentrations (Fig. 5C and Supporting Information Table S12). This was further supported by a toxicity evaluation in Kunming mice administered two representative glycoforms, **Ins6** or **Ins7**, at a dose of 1.2 units/kg daily for 14 days. Histological analysis of liver and kidney tissues revealed normal cellular architecture with no signs of necrosis, and H&E staining showed no abnormal cell morphology (Supporting Information Fig. S27). No significant differences were observed compared with saline-treated mice, confirming the absence of organ toxicity following repeated administration of glycosylated insulins.

3. Conclusions

To address the highly challenging issue of determining the impact and potential therapeutic benefits of naturally occurring *O*-linked glycosylation, we developed an enabling approach for the synthesis of insulin glycoforms bearing complex glycans. For the first time, we revealed that maintaining the fidelity of disulfide bonds while fully deprotecting different *O*-linked glycans requires distinct synthetic strategies for glycoproteins like human insulin. We discovered that glycoforms without sialic acid are best prepared using strategy I (Fig. 3), where the protecting groups on glycans are removed as the final step. For glycoforms with sialic acid, strategy II is essential, requiring glycan deprotection during the preparation of individual chains. Further refinement of strategy II was necessary based on the glycosidic bond linking the sialic acid: strategy IIa for 2,6-linked sialic acid, requiring one-step deprotection, and the more complex strategy IIb for 2,3-linked sialic acid, necessitating a two-step deprotection process (Fig. 3). Together, these strategies established a robust platform enabling the synthesis of **Ins1−7**, a series of insulin glycoforms with systematically varied glycan structures.

With this previously unattainable isoform series in hand, we were able, for the first time, to directly investigate the impact of natural *O*-glycosylation on properties and functions related to the therapeutic use of insulin. Our results revealed that natural *O*-glycosylation suppresses insulin self-association and enhances its solubility to an unexpected level, outperforming the amino acid substitutions in commercially available insulin analogs such as Aspart, Lispro, and Glulisine. These significant advantages, combined with preserved core biological functions, suggest that natural *O*-linked glycosylation holds great promise for the future development of new insulin drugs, potentially addressing two unmet medical needs. First, the markedly lower self-association propensity compared to current rapid-acting insulins may enable formulations with significantly faster onset of action. Second, the substantially higher solubility of glycosylated insulins compared to existing insulin drugs could permit development of more

concentrated formulations to benefit patients with severe insulin resistance.

Overall, we have provided a foundational framework for the total synthesis of proteins with naturally occurring *O*-glycans, highlighting the importance of having a complete repertoire of glycoforms to elucidate the therapeutic potential of *O*-glycosylation in proteins like insulin. This achievement has the potential to spur further investigations into other naturally occurring *O*-linked glycoproteins and drive advancements in utilizing natural glycosylation for therapeutic enhancements of insulin and other proteins and peptides.

4. Experimental

4.1. Synthesis and characterization of insulin glycoforms

Detailed procedures for the synthesis of insulin glycoforms, Glu-C digestion, and CD characterization are provided in the Supporting Information

4.2. AUC characterization

Insulin variants were dissolved in different buffers to final concentrations of 60 or 600 μmol/L. AUC experiments were performed using a Beckman Coulter Optima AUC equipped with an An-60 Ti analytical 4-place titanium rotor at 20 °C and 60k rpm (262,000×*g*). Each sample was analyzed alongside a corresponding buffer reference. Data were collected using Rayleigh interference detection with a 1024 × 1088 CCD camera, analyzed with the SEDFIT program, and nonlinearly fitted with a Gaussian function in Origin software. The area under each of the fit curves was used to approximate the abundance of each species.

4.3. Solubility determination

To 200 μg of lyophilized human insulin, 10 μL of PBS was added. The resulting mixture was vortexed for 2 min and centrifuged at 2680×*g* for 2 min. The supernatant was collected and diluted for UPLC analysis to obtain the UV peak area, and its concentration was interpolated from the standard curve. The solubility of human insulin in PBS was reported as the mean of three independent measurements. All other solubilities obtained in PBS or 1 mol/L PB buffer (pH 7.4) were determined in a similar way.

4.4. Proteolytic stability determination

Each insulin variant (0.5 g/L, 80 μL) in digestion buffer (100 mmol/L Tris, 1.0 mmol/L $CaCl_2$, pH 8.0) was pre-incubated at 37 °C for 15 min. Chymotrypsin (0.25 g/L, 4 μL) was added, and digestion proceeded at 37 °C for 60 min. Aliquots were withdrawn at specific intervals for UPLC and MS analyses. The remaining intact insulin variant was quantified based on UV peak area. Data were fitted according to Eq. (1):

$$y = A1 \cdot e^{\frac{x}{t1}} + y0 \quad (1)$$

where y = UV peak area, x = digestion time, and $t1$ = decay constant. Half-lives (H) were calculated as Eq. (2):

$$H = t1 \cdot ln\,2 \quad (2)$$

4.5. Long-term storage stability determination

Solutions of insulin, **Ins6**, and **Ins7** were prepared at concentrations of 60 μmol/L in PBS or 600 μmol/L in a rapid-acting insulin formulation buffer (16 g/L glycerol, 1.50 g/L phenol, 1.72 g/L meta-cresol, 19.6 mg/L Zn^{2+}, 1.25 g/L $Na_2HPO_4 \cdot 2H_2O$, 0.58 g/L NaCl, pH 7.4). The solutions were incubated at room temperature for 21 days. Samples were withdrawn at different time intervals for UPLC and MS analyses. The remaining concentration of intact insulin variant was quantified based on UV peak area.

4.6. Glucose uptake regulation determination

C57BL/6N mice were fasted for 5 h, then injected intraperitoneally with streptozotocin (160 mg/kg in 0.1 mol/L citrate buffer, pH 4.5). After 2 h, food was returned. Hyperglycemia (fasted blood glucose ≥16.7 mmol/L) was confirmed after 7 days. For functional studies, mice were fasted for 6 h, then administered insulin variants (0.75 units/kg in saline) subcutaneously. Blood glucose levels were measured at 15, 30, 45, 60, 90, and 120 min post-injection. Data were analyzed using Origin software, with results calculated from seven replicates per group.

All animals in our study received humane care, and all experimental protocols were approved by the Animal Care & Welfare Committee of Institute of Materia Medica, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China (approval number: 00003455).

4.7. AKT phosphorylation immunoblot analysis

HepG2 cells were cultured in DMEM supplemented with 10% FBS and 1% penicillin/streptomycin under a humidified atmosphere of 95% air and 5% CO_2 at 37 °C. Cells were plated in 6-well plates at a density of 1.0×10^6 live cells per well in 2 mL of medium. After 6 h, the medium was replaced with a serum-free culture medium, and cells were incubated for an additional 12 h. Subsequently, the cells were treated with insulin variants at specific concentrations for predetermined durations. Cells were then washed with ice-cold PBS and lysed with an ice-cold lysis buffer containing phosphatase and protease inhibitors for 20 min. Lysates were collected and centrifuged at 9400×*g* for 15 min. The supernatant was mixed with 1/4 volume of SDS-PAGE loading buffer, incubated at 95 °C for 10 min, and then immediately cooled on ice. Protein samples were resolved by SDS-PAGE and transferred to a PVDF membrane.

Membranes were blocked with 5% skimmed milk in TBST. Primary antibody incubation was performed overnight at 4 °C in a commercial primary antibody dilution buffer. Following six 5-min TBST washes, membranes were incubated with an HRP-conjugated secondary antibody in a commercial HRP-antibody dilution buffer for 1 h. After another six 5-min TBST washes, membranes were treated with Super ECL Detection Reagent for 3 min. Signals were acquired using the MiniChemi chemiluminescence imaging system, and immunoblot quantification was performed using Fiji software.

4.8. Toxicity in cells

HepG2 and HEK293 cells were cultured as above, seeded in 96-well plates at a density of 6000 live cells per well in 100 μL of

culture medium on the day before the assay. The culture medium was then replaced with 100 μL of fresh medium containing the desired concentration of insulin variants or control medium. The cells were incubated for 48 h, followed by the addition of the CCK-8 reagent. Absorbance was measured at 450 nm after 2 h using the Multiskan™ FC microplate photometer. Data analysis was conducted using Origin software, and results were calculated based on three replicates per group.

4.9. Toxicity in mice

Kunming mice were subcutaneously injected daily with 1.2 units/kg of insulin, **Ins6**, or **Ins7** in saline, or saline alone, for 14 consecutive days. Following euthanasia, liver and kidney tissues were collected for H&E staining.

Acknowledgments

We thank Thomas R. Cech, Virginia W. Cornish, Rudolph L. Leibel, and Li Qiang for their invaluable suggestions and advice during the preparation of this manuscript, and Tarek Sammakia and Xuefei Huang for their invaluable assistance and support. This work was supported by the CAMS Innovation Fund for Medical Sciences (CIFMS, 2025-I2M-KJ-014 and 2021-I2M-1-026, China), the National Key Research and Development Program of China (Grant No. 2018YFE0111400), the National Natural Science Foundation of China (Training Program of the Major Research Plan, Grant No. 91853120), the National Major Scientific and Technological Special Project of China (Grant Nos. 2018ZX09711001-005 and 2018ZX09711001-013), and the NIH Research Project Grant Program (Grant No. R01 EB025892, China).

Author contributions

Zhongping Tan initiated and supervised the project; Yaohao Li, Wenqiang Liu, and Yajing Zhang performed the chemical synthesis; Yaohao Li, Wenqiang Liu, Ruihan Wang, and Shiying Shang carried out the AUC experiments; Yaohao Li conducted the solubility studies; Yaohao Li and Jinyuan Gong performed the proteolytic stability studies; Yaohao Li, Dan Liu, and Xin Li conducted the functional studies; Yaohao Li, Shiying Shang, and Zhongping Tan wrote the manuscript.

Conflicts of interest

Two patent applications related to this work have been filed by the Institute of Materia Medica, Chinese Academy of Medical Sciences and Peking Union Medical College.

Appendix A. Supporting information

Supporting information to this article can be found online at https://doi.org/10.1016/j.apsb.2025.08.005.

References

1. Flier JS, Kahn CR. Insulin: a pacesetter for the shape of modern biomedical science and the Nobel prize. *Mol Metabol* 2021;**52**:101194.
2. Harrison VS, Khan MH, Chamberlain CE, Harlan DM. The noble and often Nobel role played by insulin-focused research in modern medicine. *Diabetes Care* 2021;**45**:23−7.
3. Zaykov AN, Mayer JP, DiMarchi RD. Pursuit of a perfect insulin. *Nat Rev Drug Discov* 2016;**15**:425−39.
4. Mathieu C, Martens PJ, Vangoitsenhoven R. One hundred years of insulin therapy. *Nat Rev Endocrinol* 2021;**17**:715−25.
5. Rosenfeld L. Insulin: discovery and controversy. *Clin Chem* 2002;**48**:2270−88.
6. Riggs AD. Making, cloning, and the expression of human insulin genes in bacteria: the path to Humulin. *Endocr Rev* 2021;**42**:374−80.
7. Cahn A, Miccoli R, Dardano A, Del Prato S. New forms of insulin and insulin therapies for the treatment of type 2 diabetes. *Lancet Diabetes Endocrinol* 2015;**3**:638−52.
8. Mathieu C, Gillard P, Benhalima K. Insulin analogues in type 1 diabetes mellitus: getting better all the time. *Nat Rev Endocrinol* 2017;**13**:385−99.
9. Yu Q, Canales A, Glover MS, Das R, Shi X, Liu Y, et al. Targeted mass spectrometry approach enabled discovery of *O*-glycosylated insulin and related signaling peptides in mouse and human pancreatic islets. *Anal Chem* 2017;**89**:9184−91.
10. Madsen TD, Hansen LH, Hintze J, Ye Z, Jebari S, Andersen DB, et al. An atlas of *O*-linked glycosylation on peptide hormones reveals diverse biological roles. *Nat Commun* 2020;**11**:4033.
11. Saito T, Itoh T. Variations and distributions of *O*-glycosidically linked sugar chains in bovine κ-casein. *J Dairy Sci* 1992;**75**:1768−74.
12. Steen PVd, Rudd PM, Dwek RA, Opdenakker G. Concepts and principles of *O*-linked glycosylation. *Crit Rev Biochem Mol Biol* 1998;**33**:151−208.
13. Runkel L, Meier W, Pepinsky RB, Karpusas M, Whitty A, Kimball K, et al. Structural and functional differences between glycosylated and non-glycosylated forms of human interferon-β (IFN-β). *Pharm Res* 1998;**15**:641−9.
14. Ioannou YA, Zeidner KM, Grace ME, Desnick RJ. Human α-galactosidase A: glycosylation site 3 is essential for enzyme solubility. *Biochem J* 1998;**332**:789−97.
15. Delorme E, Lorenzini T, Giffin J, Martin F, Jacobsen F, Boone T, et al. Role of glycosylation on the secretion and biological activity of erythropoietin. *Biochemistry* 1992;**31**:9871−6.
16. Guan X, Chaffey PK, Wei X, Gulbranson DR, Ruan Y, Wang X, et al. Chemically precise glycoengineering improves human insulin. *ACS Chem Biol* 2018;**13**:73−81.
17. Hossain MA, Okamoto R, Karas JA, Praveen P, Liu MJ, Forbes BE, et al. Total chemical synthesis of a nonfibrillating human glycoinsulin. *J Am Chem Soc* 2020;**142**:1164−9.
18. Kabotso DEK, Smiley D, Mayer JP, Gelfanov VM, Perez Tilve D, DiMarchi RD, et al. Addition of sialic acid to insulin confers superior physical properties and bioequivalence. *J Med Chem* 2020;**63**:6134−43.
19. Zhao J, Ye F, Huang P, Wang P. Recent advances in chemical synthesis of *O*-linked glycopeptides and glycoproteins: an advanced synthetic tool for exploring the biological realm. *Curr Opin Chem Biol* 2023;**77**:102405.
20. Karas JA, Wade JD, Hossain MA. The chemical synthesis of insulin: an enduring challenge. *Chem Rev* 2021;**121**:4531−60.
21. Gamblin DP, Scanlan EM, Davis BG. Glycoprotein synthesis: an update. *Chem Rev* 2009;**109**:131−63.
22. Liu F, Luo EY, Flora DB, Mezo AR. A synthetic route to human insulin using isoacyl peptides. *Angew Chem Int Ed* 2014;**53**:3983−7.
23. Postma TM, Giraud M, Albericio F. Trimethoxyphenylthio as a highly labile replacement for *tert*-butylthio cysteine protection in Fmoc solid phase synthesis. *Org Lett* 2012;**14**:5468−71.
24. Dziadek S, Kunz H. Synthesis of a tumor-associated 2,3-sialyl-T glycododecapeptide antigen from the tandem repeat region of the mucin MUC1. *Synlett* 2003;**2003**:1623−6.
25. Schwarz JB, Kuduk SD, Chen XT, Sames D, Glunz PW, Danishefsky SJ. A broadly applicable method for the efficient

synthesis of α-*O*-linked glycopeptides and clustered sialic acid residues. *J Am Chem Soc* 1999;**121**:2662−73.
26. Bray BL. Large-scale manufacture of peptide therapeutics by chemical synthesis. *Nat Rev Drug Discov* 2003;**2**:587−93.
27. Dunn MF. Zinc-ligand interactions modulate assembly and stability of the insulin hexamer—a review. *Biometals* 2005;**18**:295−303.
28. Home PD. The pharmacokinetics and pharmacodynamics of rapid-acting insulin analogues and their clinical consequences. *Diabetes Obes Metabol* 2012;**14**:780−8.
29. Bolli GB, Porcellati F, Lucidi P, Fanelli CG, Owens DR. One-hundred year evolution of prandial insulin preparations: from animal pancreas extracts to rapid-acting analogs. *Metabolism* 2022;**126**:154935.
30. Pohl R, Hauser R, Li M, De Souza E, Feldstein R, Seibert R, et al. Ultra-rapid absorption of recombinant human insulin induced by zinc chelation and surface charge masking. *J Diabetes Sci Technol* 2012;**6**: 755−63.
31. Cunha JP. "NOVOLOG" (RxList, 2023). https://www.rxlist.com/novolog-drug.htm.
32. Solá RJ, Griebenow K. Effects of glycosylation on the stability of protein pharmaceuticals. *J Pharmacol Sci* 2009;**98**:1223−45.
33. Brusko CSA, Taylor AD, Jackson JA, Bunck MC. Clinical challenges with concentrated insulins: setting the record straight. *Diabetes Spectr* 2017;**30**:229−32.
34. Ferreira J, Castro F. Advances in protein solubility and thermodynamics: quantification, instrumentation, and perspectives. *CrystEngComm* 2023;**25**:6388−404.
35. Wiendahl M, Volker C, Husemann I, Krarup J, Staby A, Scholl S, et al. A novel method to evaluate protein solubility using a high throughput screening approach. *Chem Eng Sci* 2009;**64**:3778−88.
36. Sinclair AM, Elliott S. Glycoengineering: the effect of glycosylation on the properties of therapeutic proteins. *J Pharmacol Sci* 2005;**94**: 1626−35.
37. Rocamora F, Peralta AG, Shin S, Sorrentino J, Wu MYM, Toth EA, et al. Glycosylation shapes the efficacy and safety of diverse protein, gene and cell therapies. *Biotechnol Adv* 2023;**67**:108206.
38. Schilling RJ, Mitra AK. Degradation of insulin by trypsin and alpha-chymotrypsin. *Pharm Res* 1991;**8**:721−7.
39. Lapolla A, Tessari P, Poli T, Valerio A, Duner E, Iori E, et al. Reduced *in vivo* biological activity of *in vitro* glycosylated insulin. *Diabetes* 1988;**37**:787−91.
40. King AJ. The use of animal models in diabetes research. *Br J Pharmacol* 2012;**166**:877−94.
41. Haeusler RA, McGraw TE, Accili D. Biochemical and cellular properties of insulin receptor signalling. *Nat Rev Mol Cell Biol* 2018; **19**:31−44.
42. Huang X, Liu G, Guo J, Su Z. The PI3K/AKT pathway in obesity and type 2 diabetes. *Int J Biol Sci* 2018;**14**:1483−96.
43. Saltiel AR. Insulin signaling in health and disease. *J Clin Investig* 2021;**131**:e142241.
44. Le TKC, Dao XD, Nguyen DV, Luu DH, Bui TMH, Le TH, et al. Insulin signaling and its application. *Front Endocrinol* 2023;**14**:1226655.
45. Kiernan K, Alwarawrah Y, Nichols AG, Danzaki K, MacIver NJ. Insulin and IGF-1 have both overlapping and distinct effects on $CD4^+$ T cell mitochondria, metabolism, and function. *Sci Rep* 2024;**14**:4331.
46. Bartels T, Wäse K, Heinrichs M, Stolte M, Roome N, Scherer P, et al. Regulatory forum opinion piece: review—toxicological pathology profile and regulatory expectations for nonclinical development of insulins and insulin analogues. *Toxicol Pathol* 2016;**44**:931−46.