Article

Not peer-reviewed version

How Structural Modifications of Insulin Icodec Contributes to Its Prolonged Duration of Action: A Structural and Biophysical Perspective

Wei Li *

Posted Date: 16 November 2023

doi: 10.20944/preprints202311.1048.v1

Keywords: Insulin icodec; Molecular design; Site-specific mutation; Structural biophysics

Article

How Structural Modifications of Insulin Icodec Contributes to Its Prolonged Duration of Action: A Structural and Biophysical Perspective

Wei Li

Contrebola Institute of Computational Interstructural Biophysics, No. 88, Fuxing East Road, Nantong City 226000, Jiangsu Province, People's Republic of China; wli148@aucklanduni.ac.nz

Abstract: Insulin icodec of Novo Nordisk is a novel long-acting insulin analogue that exhibits an extended duration of action, providing a promising treatment option for individuals with diabetes. It has been reported that the incorporation of fatty acid moieties into insulin icodec plays a crucial role in its prolonged action, as these fatty acid chains facilitate the formation of stable hexameric structures, thereby delaying insulin absorption and promoting sustained release. Yet, the underlying biophysics still is elusive of the roles of the three site-specific mutations (Y14A_E, Y37B_H, F46B_H) of insulin icodec in its prolonged activity. Thus, through a comprehensive structural and biophysical analysis of the insulin (both native and icodec) structures bound to its receptor, this article delves deep into the biophysics underlying the molecular design of insulin icodec, and identified a delicate biophysical mechanism through which two missense mutations of insulin icodec (Y37B_H and F46B_H) contribute to its prolonged duration of action. Overall, this structural and biophysical investigation provides valuable insights into the mechanisms underlying the relationship between three site-specific mutations and prolonged duration of action of insulin icodec, while understanding these modifications at a structural and biophysical level can aid in the rational design of future long-acting insulin analogues, offering further enhanced therapeutic options for diabetic patients.

Keywords: insulin icodec; molecular design; site-specific mutation; structural biophysics

1. Introduction

Ligand-receptor binding affinity is an essential parameter in both computational and experimental drug discovery & design [1]. Thanks to the continued development of experimental structural biology and the fifty-three-year old Protein Data Bank (PDB) [2,3], a high-throughput comprehensive structural biophysical analysis becomes possible [4] for specific ligand-receptor complex structures deposited in PDB, such that our understanding of the structural and biophysical basis of their interfacial stability is able to help us modify the binding affinity of certain drug target and its interacting partners.

Take insulin icodec for example [5–8]. Insulin icodec of Novo Nordisk is a long-acting insulin analogue for better management of blood sugar levels in people with diabetes [9–11]. Insulin icodec is designed to provide a steady release of insulin throughout the day, mimicking the natural insulin production in the body [12–14]. Insulin icodec is typically administered through injection once a week, which helps lower blood sugar levels by allowing glucose to enter the body's cells, where it is subsequently used for energy production [15,16]. Moreover, insulin icodec has a distinct pharmacokinetic profile compared to other long-acting insulin analogs. It exhibits a long duration of action, with a half-life of approximately 196 hours, leading to improved glycemic control and reduced hypoglycemia risk [17,18]. Interestingly, Icosema (of Novo Nordisk, too) [19–22] represents a combination medication that consists of insulin icodec and semaglutide, which belongs to a class of medications called glucagon-like peptide-1 (GLP-1) receptor agonists and helps regulate blood sugar levels by stimulating the release of insulin, reducing the production of glucagon (a hormone that increases blood sugar levels), and slowing down the digestion process for weight reduction [23–25].

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 16 November 2023 doi:10.20944/preprints202311.1048.v1

Given the two-dimensional structures of native human insulin and insulin icodec as shown in Figure 1, this article summarizes below a list of structural modifications of insulin icodec compared to native human insulin [26–28]:

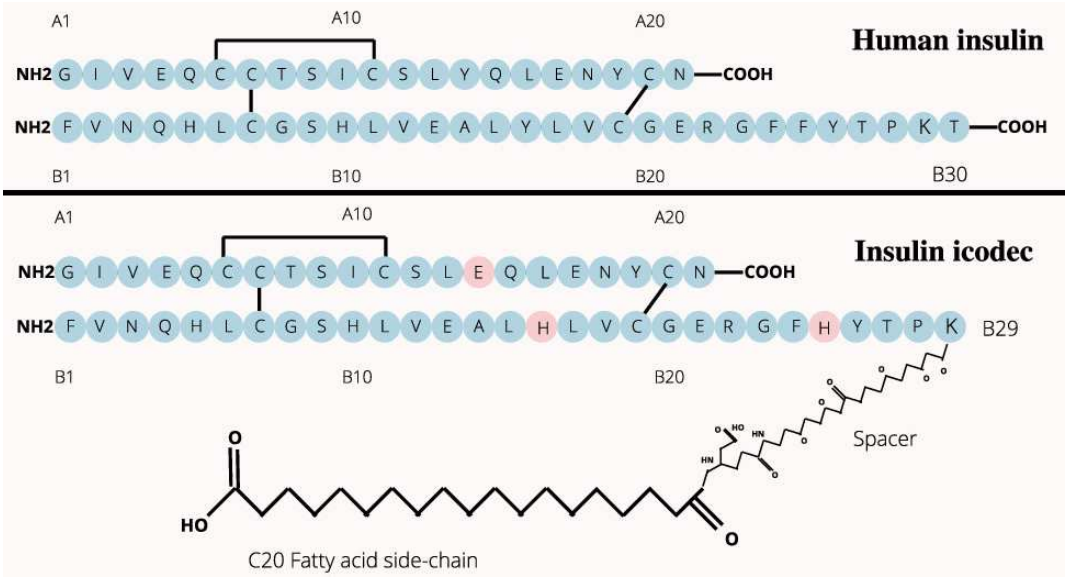

Figure 1. A brief illustration of the two-dimensional structures of native human insulin and insulin icodec of Novo Nordisk [29–31]. In this figure, the amino acid residues with pink backgrounds represents the positions of the three site-specific mutations (Y14A_E, Y37B_H, F46B_H) of insulin icodec.

1. addition of a fatty acid (K50B_C20, Figures 1 and 2);
2. deletion of threonine at position B30 (T51B_del, Figures 1 and 2);
3. a site-specific missense mutation, Y14A_E (Figures 1 and 2);
4. a site-specific missense mutation, Y37B_H (Figures 1 and 2);
5. a site-specific missense mutation, F46B_H (Figures 1 and 2).

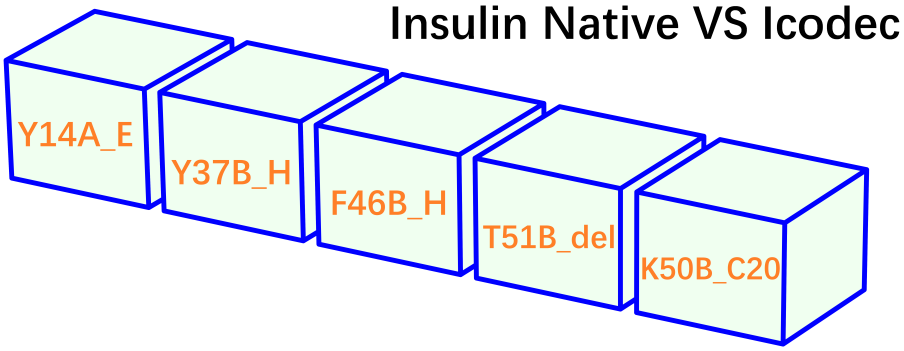

Figure 2. A summary of the structural modification of insulin icodec in comparison to native human insulin. In this figure, Y14A_E (i.e., replacement of Tyr14 (Y14) at position A14 (position 14 of chain A) by a histidine), Y37B_H (i.e., replacement of Tyr16 (Y16) at position B16 by a histidine) and F46B_H (i.e., replacement of Phe25 (F25) at position B25 by a histidine) represent three site-specific missense mutations of insulin icodec, T51B_del represents deletion of Thr30 (T30) at position B30, while K50B_C20 represents the addition of a 20-carbon fatty acid to the lysine amino acid (K50B) at position B29 [32–35].

 doi:10.20944/preprints202311.1048.v1

Overall, these structural modifications (Figure 2) in insulin icodec result in a more stable and longer-acting insulin analogue compared to regular insulin, providing a more consistent and sustained blood glucose-lowering effect:

1. insulin icodec is able to form aggregates or clusters at the subcutaneous injection site, gradually releasing into the bloodstream over an extended period.
2. insulin icodec undergoes structural modifications that increases its stability and solubility and preventing enzyme-mediated degradation and rapid clearance [36].
3. insulin icodec is conjugated with a fatty acid at position B30 (K50B_C20, Figure 2). After injection, the fatty acid chain in insulin icodec interacts with albumin in the subcutaneous tissue, forming reversible albumin-insulin complexes. These complexes act as a reservoir, gradually releasing insulin icodec into the bloodstream, increasing its fat solubility and allowing it to bind to fatty acid-binding proteins, forming a depot of the insulin icodec reversibly bound to albumin.
4. the incorporation of fatty acid chains facilitate the formation of stable hexameric structures, thereby delaying insulin absorption and promoting sustained release.

With respect to the addition of a fatty acid (K50B_C20, Figure 2), it is conceivable that the deletion of threonine at position B30 (T51B_del, Figure 2) is for K50B_C20 to take place easier and more efficiently than without the deletion of threonine at position B30. Nonetheless, the underlying biophysics remains elusive of the roles of the three site-specific mutations (Y14A_E, Y37B_H, F46B_H) of insulin icodec in its prolonged activity. Therefore, through a comprehensive structural and biophysical analysis of the insulin (both native and icodec) structures bound to its receptor, this article delves deep into the biophysics underlying the molecular design of insulin icodec, and identified a delicate biophysical mechanism through which two missense mutations of insulin icodec (Y37B_H and F46B_H) contribute to its prolonged duration of action.

2. Materials and Methods

2.1. A summary of insulin receptor-related structures in PDB

As listed in Table 1, as of November 16, 2023, there are a variety of experimental complex structures of insulin (analogues) bound to its receptor (IR), such as PDB entries 7PG3 (insulin receptor bound to 3 insulins), 7PG4 (insulin receptor bound to 2 insulins), 6SOF (insulin receptor bound to 4 insulins).

Table 1. Experimentally determined IR-related structures (released newest from oldest) in the Protein Data Bank (PDB [2]) as of November 16, 2023, QUERY code: UniProt Molecule Name = "Insulin receptor".

PDB ID	Structure Title (release date from newest to oldest)
8DWN	Crystal structure of bis-phosphorylated insulin receptor kinase domain
7YQ3	human insulin receptor bound with A43 DNA aptamer and insulin
7YQ4	human insulin receptor bound with A62 DNA aptamer and insulin - locally refined
7YQ5	human insulin receptor bound with A62 DNA aptamer and insulin
7YQ6	human insulin receptor bound with A62 DNA aptamer
8EYX	Cryo-EM structure of 4 insulins bound full-length mouse IR mutant with physically decoupled alpha CTs (C684S/C685S/C687S; denoted as IR-3CS) Asymmetric conformation 1
8EYY	Cryo-EM structure of 4 insulins bound full-length mouse IR mutant with physically decoupled alpha CTs (C684S/C685S/C687S, denoted as IR-3CS) Asymmetric conformation 2
8EZ0	Cryo-EM structure of 4 insulins bound full-length mouse IR mutant with physically decoupled alpha CTs (C684S/C685S/C687S; denoted as IR-3CS) Symmetric conformation
8GUY	human insulin receptor bound with two insulin molecules
7U6D	Head region of insulin receptor ectodomain (A-isoform) bound to the non-insulin agonist IM459
7U6E	Head region of insulin receptor ectodomain (A-isoform) bound to the non-insulin agonist IM462
7PHT	Structure of Insulin receptor's transmembrane domain
8DTL	Cryo-EM structure of insulin receptor (IR) bound with S597 peptide

 doi:10.20944/preprints202311.1048.v1

Table 1. *Cont.*

PDB ID	Structure Title (release date from newest to oldest)
8DTM	Cryo-EM structure of insulin receptor (IR) bound with S597 component 2
7S0Q	Head region of a complex of IGF-I with the ectodomain of a hybrid insulin receptor / type 1 insulin-like growth factor receptor
7S8V	Leg region of a complex of IGF-I with the ectodomain of a hybrid insulin receptor / type 1 insulin-like growth factor receptor
7SL1	Full-length insulin receptor bound with site 1 binding deficient mutant insulin (A-V3E)
7SL2	Full-length insulin receptor bound with site 2 binding deficient mutant insulin (A-L13R) – asymmetric conformation
7SL3	Full-length insulin receptor bound with site 2 binding deficient mutant insulin (A-L13R) – symmetric conformation
7SL4	Full-length insulin receptor bound with site 2 binding deficient mutant insulin (B-L17R) – asymmetric conformation
7SL6	Full-length insulin receptor bound with site 2 binding deficient mutant insulin (B-L17R) – symmetric conformation
7SL7	Full-length insulin receptor bound with both site 1 binding deficient mutant insulin (A-V3E) and site 2 binding deficient mutant insulin (A-L13R)
7STH	Full-length insulin receptor bound with unsaturated insulin WT (2 insulin bound) symmetric conformation
7STI	Full-length insulin receptor bound with unsaturated insulin WT (1 insulin bound) asymmetric conformation
7STJ	Full-length insulin receptor bound with unsaturated insulin WT (2 insulins bound) asymmetric conformation (Conformation 1)
7STK	Full-length insulin receptor bound with unsaturated insulin WT (2 insulins bound) asymmetric conformation (Conformation 2)
7MQO	The insulin receptor ectodomain in complex with a venom hybrid insulin analogue - "head" region
7MQR	The insulin receptor ectodomain in complex with four venom hybrid insulins - symmetric conformation
7MQS	The insulin receptor ectodomain in complex with three venom hybrid insulin molecules - asymmetric conformation
7MD4	Insulin receptor ectodomain dimer complexed with two IRPA-3 partial agonists
7MD5	Insulin receptor ectodomain dimer complexed with two IRPA-9 partial agonists
7PG0	Low resolution Cryo-EM structure of full-length insulin receptor bound to 3 insulin with visible ddm micelle, conf 1
7PG2	Low resolution Cryo-EM structure of full-length insulin receptor bound to 3 insulin, conf 1
7PG3	Low resolution Cryo-EM structure of the full-length insulin receptor bound to 3 insulin, conf 2
7PG4	Low resolution Cryo-EM structure of the full-length insulin receptor bound to 2 insulin, conf 3
7QID	tentative model of the human insulin receptor ectodomain bound by three insulin
7KD6	Insulin Receptor L1-CR plus alphaCT fragment in co-complex with Fv 83-7 and single-chain insulin SCI-b
7BW7	Cryo-EM Structure for the Ectodomain of the Full-length Human Insulin Receptor in Complex with 1 Insulin.
7BW8	Cryo-EM Structure for the Insulin Binding Region in the Ectodomain of the Full-length Human Insulin Receptor in Complex with 1 Insulin
7BWA	Cryo-EM Structure for the Ectodomain of the Full-length Human Insulin Receptor in Complex with 2 Insulin
6VEP	Human insulin in complex with the human insulin microreceptor in turn in complex with Fv 83-7
6VEQ	Con-Ins G1 in complex with the human insulin microreceptor in turn in complex with Fv 83-7
6SOF	human insulin receptor ectodomain bound by 4 insulin
6PXV	Cryo-EM structure of full-length insulin receptor bound to 4 insulin. 3D refinement was focused on the extracellular region.
6PXW	Cryo-EM structure of full-length insulin receptor bound to 4 insulin. 3D refinement was focused on the top part of the receptor complex.
6HN4	Leucine-zippered human insulin receptor ectodomain with single bound insulin - "lower" membrane-proximal part
6HN5	Leucine-zippered human insulin receptor ectodomain with single bound insulin - "upper" membrane-distal part
6CE7	Insulin Receptor ectodomain in complex with one insulin molecule
6CE9	Insulin Receptor ectodomain in complex with two insulin molecules

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 16 November 2023 doi:10.20944/preprints202311.1048.v1

Table 1. *Cont.*

PDB ID	Structure Title (release date from newest to oldest)
6CEB	Insulin Receptor ectodomain in complex with two insulin molecules - C1 symmetry
5U1M	Structure of the IRS-1 PTB Domain Bound to the Juxtamembrane Region of the Insulin Receptor
5KQV	Insulin receptor ectodomain construct comprising domains L1,CR,L2, FnIII-1 and alphaCT peptide in complex with bovine insulin and FAB 83-14 (REVISED STRUCTURE)
5TQ1	Phospholipase C gamma-1 C-terminal SH2 domain bound to a phosphopeptide derived from the insulin receptor
5J3H	Human insulin receptor domains L1-CR in complex with peptide S519C16 and 83-7 Fv
5HHW	Crystal structure of insulin receptor kinase domain in complex with cis-(R)-7-(3-(azetidin-1-ylmethyl)cyclobutyl)-5-(3-((tetrahydro-2H-pyran-2-yl)methoxy)phenyl)-7H -pyrrolo[2,3-d]pyrimidin-4-amine.
4ZXB	Structure of the human insulin receptor ectodomain, IRDeltabeta construct, in complex with four Fab molecules
5E1S	The Crystal structure of INSR Tyrosine Kinase in complex with the Inhibitor BI 885578
4XSS	Insulin-like growth factor I in complex with site 1 of a hybrid insulin receptor / Type 1 insulin-like growth factor receptor
4XST	Structure of the endoglycosidase-H treated L1-CR domains of the human insulin receptor in complex with residues 697-719 of the human insulin receptor (A-isoform)
4XLV	Crystal structure of the activated insulin receptor tyrosine kinase dimer
4OGA	Insulin in complex with Site 1 of the human insulin receptor
2MFR	Solution structure of the transmembrane domain of the insulin receptor in micelles
4IBM	Crystal structure of insulin receptor kinase domain in complex with an inhibitor Irfin-1
3W11	Insulin receptor ectodomain construct comprising domains L1-CR in complex with human insulin, Alpha-CT peptide(704-719) and FAB 83-7
3W12	Insulin receptor ectodomain construct comprising domains L1-CR in complex with high-affinity insulin analogue [D-PRO-B26]-DTI-NH2, alpha-CT peptide(704-719) and FAB 83-7
3W13	Insulin receptor ectodomain construct comprising domains L1-CR in complex with high-affinity insulin analogue [D-PRO-B26]-DTI-NH2, alphact peptide(693-719) and FAB 83-7
3ETA	Kinase domain of insulin receptor complexed with a pyrrolo pyridine inhibitor
3EKN	Insulin receptor kinase complexed with an inhibitor
3EKK	Insulin receptor kinase complexed with an inhibitor
2Z8C	Phosphorylated insulin receptor tyrosine kinase in complex with (4-[5-carbamoyl-4-(3-methylanilino)pyrimidin-2-yl]aminophenyl)acetic acid
3BU3	Crystal structure of the insulin receptor kinase in complex with IRS2 KRLB peptide
3BU5	Crystal structure of the insulin receptor kinase in complex with IRS2 KRLB peptide and ATP
3BU6	Crystal structure of the insulin receptor kinase in complex with IRS2 KRLB phosphopeptide
2HR7	Insulin receptor (domains 1-3)
2B4S	Crystal structure of a complex between PTP1B and the insulin receptor tyrosine kinase
2AUH	Crystal structure of the Grb14 BPS region in complex with the insulin receptor tyrosine kinase
1RQQ	Crystal Structure of the Insulin Receptor Kinase in Complex with the SH2 Domain of APS
1LK2	1.35A crystal structure of H-2Kb complexed with the GNYSFYAL peptide
1P14	Crystal structure of a catalytic-loop mutant of the insulin receptor tyrosine kinase
1I44	CRYSTALLOGRAPHIC STUDIES OF AN ACTIVATION LOOP MUTANT OF THE INSULIN RECEPTOR TYROSINE KINASE
1GAG	CRYSTAL STRUCTURE OF THE INSULIN RECEPTOR KINASE IN COMPLEX WITH A BISUBSTRATE INHIBITOR
1IR3	PHOSPHORYLATED INSULIN RECEPTOR TYROSINE KINASE IN COMPLEX WITH PEPTIDE SUBSTRATE AND ATP ANALOG
1IRK	CRYSTAL STRUCTURE OF THE TYROSINE KINASE DOMAIN OF THE HUMAN INSULIN RECEPTOR

Among the insulin receptor-related structures listed in Table 1, PDB entry 6SOF [37] is the only experimental complex structure of insulin bound to IR, where all four distinct binding sites of the IR dimer are saturated by four insulin molecules. Therefore, PDB entry 6SOF [37] is chosen here as the structural template for subsequent structural modelling of insulin icodec bound to IR.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 16 November 2023 doi:10.20944/preprints202311.1048.v1

2.2. Construction of a complex structural model of insulin icodec bound to IR

Towards the construction of a complex structural model of insulin icodec bound to IR, first, the amino acid sequences of the two chains of native human insulin and IR (according to PDB entry 6SOF [37], supplementary file **nati.pdb**) are listed in italics in fasta format as below,

>InsulinA

GIVEQCCTSICSLYQLENYCN

>InsulinB

FVNQHLCGSHLVEALYLVCGERGFFYTPKT

>IRpartOne

HLYPGEVCPGMDIRNNLTRLHELENCSVIEGHLQILLMFKTRPEDFRDLSFPKLIMITDYLLLFRV
YGLESLKDLFPNLTVIRGSRLFFNYALVIFEMVHLKELGLYNLMNITRGSVRIEKNNELCYLATIDWSRI
LDSVEDNYIVLNKDDNEECGDICPGTAKGKTNCPATVINGQFVERCWTHSHCQKVCPTICKSHGCT
AEGLCCHSECLGNCSQPDDPTKCVACRNFYLDGRCVETCPPPYYHFQDWRCVNFSFCQDLHHKCK
NSRRQGCHQYVIHNNKCIPECPSGYTMNSSNLLCTPCLGPCPKVCHLLEGEKTIDSVTSAQELRGCTV
INGSLIINIRGGNNLAAELEANLGLIEEISGYLKIRRSYALVSLSFFRKLRLIRGETLEIGNYSFYALDNQN
LRQLWDWSKHNLTITQGKLFFHYNPKLCLSEIHKMEEVSGTKGRQERNDIALKTNGDQASCENELLK
FSYIRTSFDKILLRWEPYWPPDFRDLLGFMLFYKEAPYQNVTEFDGQDACGSNSWTVVDIDPPLRSN
DPKSQNHPGWLMRGLKPWTQYAIFVKTLVTFSDERRTYGAKSDIIYVQTDATNPSVPLDPISVSNSSS
QIILKWKPPSDPNGNITHYLVFWERQAEDSELFELDYCLKGLKLPSRTWSPPFESEDSQKHNQSEYED
SAGECCSCPKTDSQILKELEESSFRKTFEDYLHNVVFVPRPS

>IRpartTwo

HRPFEKVVNKESLVISGLRHFTGYRIELQACNQDTPEERCSVAAYVSARTMPEAKADDIVGPVT
HEIFENNVVHLMWQEPKEPNGLIVLYEVSYRRYGDEELHLCVSRKHFALERGCRLRGLSPGNYSVRI
RATSLAGNGSWTEPTYFYVTDYLDVPSNIAK

Second, the amino acid sequences of the two chains of insulin icodec and IR are listed in italics in fasta format as below,

>InsulinIcodecA

GIVEQCCTSICSLEQLENYCN

>InsulinIcodecB

FVNQHLCGSHLVEALHLVCGERGFFHTPKT

>IRpartOne

HLYPGEVCPGMDIRNNLTRLHELENCSVIEGHLQILLMFKTRPEDFRDLSFPKLIMITDYLLLFRV
YGLESLKDLFPNLTVIRGSRLFFNYALVIFEMVHLKELGLYNLMNITRGSVRIEKNNELCYLATIDWSRI
LDSVEDNYIVLNKDDNEECGDICPGTAKGKTNCPATVINGQFVERCWTHSHCQKVCPTICKSHGCT
AEGLCCHSECLGNCSQPDDPTKCVACRNFYLDGRCVETCPPPYYHFQDWRCVNFSFCQDLHHKCK
NSRRQGCHQYVIHNNKCIPECPSGYTMNSSNLLCTPCLGPCPKVCHLLEGEKTIDSVTSAQELRGCTV
INGSLIINIRGGNNLAAELEANLGLIEEISGYLKIRRSYALVSLSFFRKLRLIRGETLEIGNYSFYALDNQN
LRQLWDWSKHNLTITQGKLFFHYNPKLCLSEIHKMEEVSGTKGRQERNDIALKTNGDQASCENELLK
FSYIRTSFDKILLRWEPYWPPDFRDLLGFMLFYKEAPYQNVTEFDGQDACGSNSWTVVDIDPPLRSN
DPKSQNHPGWLMRGLKPWTQYAIFVKTLVTFSDERRTYGAKSDIIYVQTDATNPSVPLDPISVSNSSS
QIILKWKPPSDPNGNITHYLVFWERQAEDSELFELDYCLKGLKLPSRTWSPPFESEDSQKHNQSEYED
SAGECCSCPKTDSQILKELEESSFRKTFEDYLHNVVFVPRPS

>IRpartTwo

HRPFEKVVNKESLVISGLRHFTGYRIELQACNQDTPEERCSVAAYVSARTMPEAKADDIVGPVT
HEIFENNVVHLMWQEPKEPNGLIVLYEVSYRRYGDEELHLCVSRKHFALERGCRLRGLSPGNYSVRI
RATSLAGNGSWTEPTYFYVTDYLDVPSNIAK

Lastly, to build a reliable structural model of the insulin icodec bound to IR, the amino acid sequences above (>InsulinIcodecA, >InsulinIcodecB, >IRpartOne and >IRpartTwo) were plugged into the SWISS-MODEL homology modeling server [38,39] and the Modeller software [40] to build a homology structural model with reasonable accuracy, supplementary file **icod.pdb**, which was

 doi:10.20944/preprints202311.1048.v1

subsequently subject to a set of comprehensive structural biophysical analysis as described previously in [4].

2.3. A comprehensive structural and biophysical analysis of insulin icodec bound to IR

As described previously in [4], a comprehensive structural and biophysical analysis was conducted for the structural model of the insulin icodec bound to IR (supplementary file **icod.pdb**). Specifically, the salt bridge analysis was conducted with an in-house python script only for titrateable residues (Asp, Glu, Lys, Arg and His), 4.0 Å was used as the cutoff distance for the two oppositely charged groups [4]. The hydrogen bond analysis was also conducted for only side chain nuclei with an in-house python script, and employed two geometric criteria: (a) a cutoff value of the angle formed by acceptor (A), donor (D) and hydrogen (H) ($\angle ADH$) of 30°; (b) a cutoff value of donor-acceptor distance at 3.0 Å. That is, a hydrogen bond is only considered to be formed if $\angle ADH$ is no larger than 30° and the donor-acceptor distance is not larger than 3.0 Å [4].

3. Results

As described above, the addition of a fatty acid to insulin icodec (K50B_C20, Figure 2) is for it to interact with albumin and form reversible albumin-insulin complexes, i.e., a reservoir for gradually releasing insulin icodec into the bloodstream. Moreover, it is highly likely that the deletion of threonine at position B30 (T51B_del, Figure 2) is for K50B_C20 to take place easier and more efficiently. As a result, the comprehensive structural and biophysical analysis here focuses on the three site-specific mutations (Y14A_E, Y37B_H, F46B_H) of insulin icodec.

Coincidentally, the three mutations (Y14A_E, Y37B_H, F46B_H) invariably fall into the same category of the substitution of an amino acid residue with an amino acid residue with a hydrophilic side chain. Briefly, with a set of comprehensive structural and biophysical analysis [4] for supplementary files **icod.pdb** and **nati.pdb**, this article here for the first time reports that

1. Y14A_E (Figures 1 and 2) is not involved in any salt bridge or hydrogen bond within insulin icodec itself.
2. Y14A_E (Figures 1 and 2) is not involved in any salt bridge or hydrogen bond at the binding interface between insulin icodec and IR.
3. Y37B_H (Figures 1 and 2) is involved in one salt bridge within insulin icodec itself (Figure 3, Tables 2 and 4).
4. Y37B_H (Figures 1 and 2) is not involved in any hydrogen bond within insulin icodec itself.
5. Y37B_H (Figures 1 and 2) is not involved in any salt bridge or hydrogen bond at the binding interface between insulin icodec and IR.
6. F46B_H (Figures 1 and 2) is not involved in any salt bridge or hydrogen bond within insulin icodec itself.
7. F46B_H (Figures 1 and 2) is not involved in any salt bridge or hydrogen bond at the binding interface between insulin icodec and IR.

Of the three mutations (Y14A_E, Y37B_H, F46B_H), a total of three electrostatic interactions (including salt bridges and hydrogen bond) were identified and listed as below in Table 2.

Table 2. Interfacial salt bridging network analysis [4] of the structural model of the insulin icodec bound to IR (supplementary files **icod.pdb** and **nati.pdb**). In this table, the residue naming scheme is **Chain ID_residue name_residue number**.

PDB file name	Residue A	Atom A	Residue B	Atom B	Distance (Å)
icod.pdb	C_LYS_921	NZ	F_GLU_1804	OE1	3.659
icod.pdb	F_HIS_1799	NE2	F_GLU_1804	OE2	3.669
nati.pdb	C_LYS_921	NZ	F_GLU_1804	OE2	3.204

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 16 November 2023 doi:10.20944/preprints202311.1048.v1

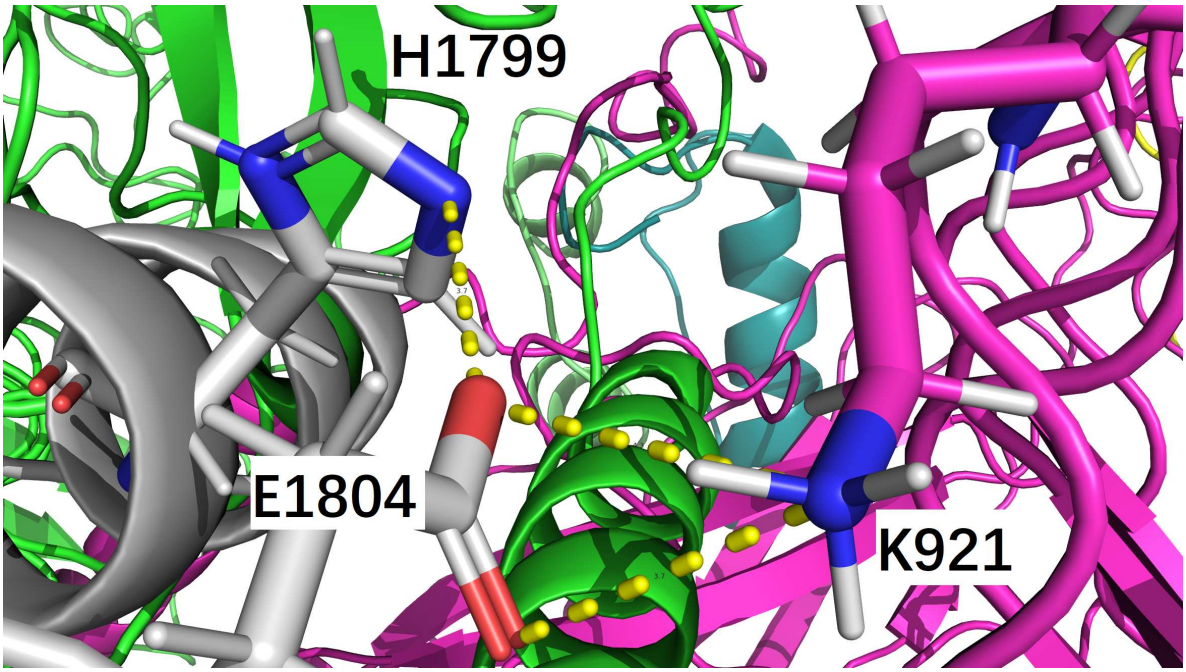

Figure 3. Glu1804 of insulin icodec forms two salt bridges, one with His1799 of insulin icodec (due to the Y37B_H substitution), and another one with Lys921 of IR at their binding interface. This figure is prepared by PyMol [41].

Specifically, before Y37B_H (Figures 1 and 2), i.e., in the case of native human insulin (Table 3), only one salt bridge (3.204 Å long) was formed between C_LYS_921 (of IR) and F_GLU_1804 (of regular insulin) at the binding interface of native human insulin bound to IR (supplementary file **nati.pdb**). Similarly, after Y37B_H (Figures 1 and 2), i.e., in the case of insulin icodec (Table 4), only one salt bridge (3.669 Å long) was formed between C_LYS_921 (of IR) and F_GLU_1804 (of insulin icodec) at the binding interface of insulin icodec bound to IR (supplementary file **icod.pdb**). Interestingly, apart from the interfacial salt bridge between C_LYS_921 (of IR) and F_GLU_1804 (of insulin icodec), there is also one intra-molecular salt bridge (3.659 Å long) between F_HIS_1799 (of insulin icodec) and F_GLU_1804 (of insulin icodec).

Table 3. Interfacial salt bridging network analysis [4] of the structural model of the insulin icodec bound to IR (supplementary file **nati.pdb**). In this table, the residue naming scheme is **Chain ID_residue name_residue number**.

PDB file name	Residue A	Atom A	Residue B	Atom B	Distance (Å)
nati.pdb	C_LYS_921	NZ	F_GLU_1804	OE2	3.204

Thus, it is obvious that in the case of native human insulin (Table 3), F_GLU_1804 (of regular insulin) focuses on binding to C_LYS_921 (of IR) alone with a 3.204 Å long interfacial salt bridge between regular insulin and IR, while in the case of insulin icodec (Table 4), F_GLU_1804 (of insulin icodec) has to bind both C_LYS_921 (of IR) and F_HIS_1799 (of insulin icodec) with two salt bridges, as listed in Table 4.

Table 4. Interfacial salt bridging network analysis [4] of the structural model of the insulin icodec bound to IR (supplementary file **icod.pdb**). In this table, the residue naming scheme is **Chain ID_residue name_residue number**.

PDB file name	Residue A	Atom A	Residue B	Atom B	Distance (Å)
icod.pdb	C_LYS_921	NZ	F_GLU_1804	OE1	3.659
icod.pdb	F_HIS_1799	NE2	F_GLU_1804	OE2	3.669

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 16 November 2023 doi:10.20944/preprints202311.1048.v1

Furthermore, the existence of F_HIS_1799 (of insulin icodec) is entirely due to the site-directed mutation Y37B_H (Figures 1 and 2). Therefore, it is entirely due to the site-directed mutation Y37B_H that the salt bridge at the binding interface between insulin and IR becomes weaker (from 3.204 Å long to 3.669 Å long), but is still not disrupted by the site-directed mutation Y37B_H, such that the binding affinity of ligand-receptor is lowered but not eliminated by the site-directed mutation Y37B_H (Figures 1 and 2), and thereby ensuring downstream signal transduction for the prolonged blood glucose-lowering effect of insulin icodec.

With respect to the site-directed mutation F46B_H (Figures 1 and 2), it is not involved in any salt bridge or hydrogen bond within insulin icodec itself, nor is it involved in any salt bridge or hydrogen bond at the binding interface between insulin icodec and IR, according comprehensive structural and biophysical analysis [4] of insulin icodec bound to IR (supplementary file **icod.pdb**). Nonetheless, a close inspection with PyMol [41] of supplementary file **icod.pdb** revealed that the positively charged side chain of His1808 of insulin icodec (due to the site-specific mutation F46B_H, Figure 2) is only 6.4 Å away (Figure 4) from the positively charged side chain of R900 (Arg900) of IR. Despite the cutoff distance (4.0 Å) as used in [4], a distance 6.4 Å is not sufficient to be defined as strong an electrostatic interaction as a salt bridge, it still is able to induce a weak electrostatic repulsive force between the positively charged side chains of His1808 of insulin icodec and that of R900 of IR, and thereby reducing the binding affinity between insulin icodec and IR. Moreover, in case the site-specific mutation is not F46B_H, Figure 2), but F46B_E, i.e., it is not a histidine, but a glutamate, it is conceivable that the distance between Glu1808 of insulin icodec and R900 of IR is shorter than 6.4 Å (Figure 4), make it highly likely for a novel stable salt bridge to be formed between E1808 (Glu1808) of insulin icodec and R900 (Arg900) of IR, according to the 4 Å distance cutoff as used in [4].

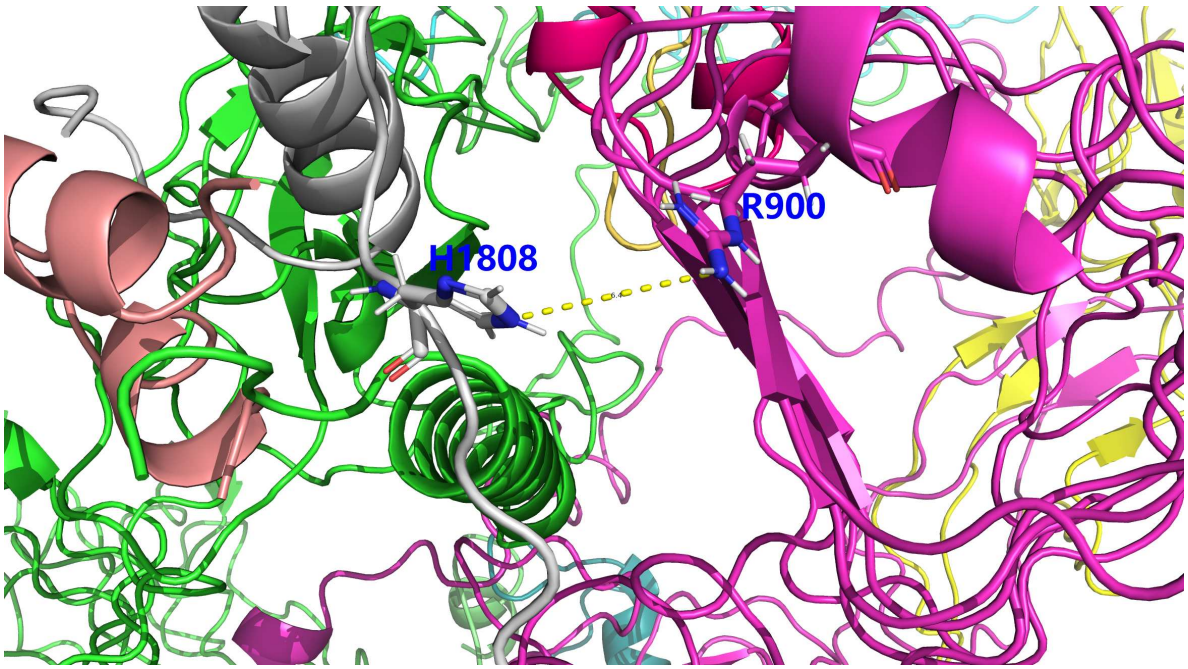

Figure 4. A 6.4 Å long electrostatic repulsive force between between H1808 of insulin icodec (His1808, due to the site-specific mutation F46B_H, Figure 2) and R900 of IR. This figure is prepared by PyMol [41] according to supplementary file **icod.pdb**.

Finally, with respect to the site-directed mutation Y14A_E (Figures 1 and 2), , it mutation does not seem to impact the binding between insulin icodec and IR, as it is a fully solvent-exposed amino acid residue, as shown by Figure 5. Thus, as of November 16, 2023 from structural and biophysical perspective, it is not clear yet why site-directed mutation Y14A_E is included in the design of insulin

 doi:10.20944/preprints202311.1048.v1

icodec. Here, this article is only able to speculate that Y14A_E is for the overall structural stability of icodec insulin for it to exert its prolonged blood glucose-lowering effect for patients with diabetes.

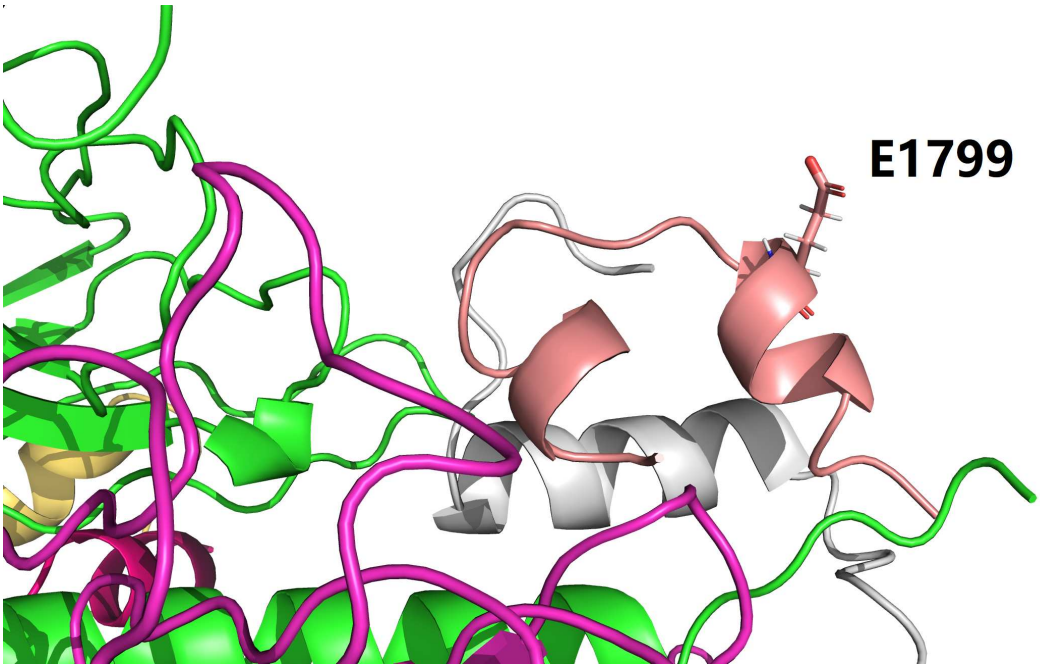

Figure 5. A fully solvent-exposed Glu14 of insulin icodec according to supplementary file **icod.pdb**. This figure is prepared by PyMol [41].

4. Conclusion and Discussion

For the first time, through a comprehensive structural and biophysical analysis of the insulin (both native and icodec) structures bound to its receptor, this article puts forward a delicate biophysical mechanism through which two missense mutations of insulin icodec (Y37B_H and F46B_H) contribute to its prolonged duration of action. In addition to the biophysics underlying the molecular design of insulin icodec, this article also argues that this structural and biophysical understanding is of help for rational design of future long-acting insulin analogues, offering further enhanced therapeutic options for diabetic patients.

5. Ethical statement

No ethical approval is required.

6. Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work, the author used OpenAI's ChatGPT in order to improve the readability of the manuscript, and to make it as concise and short as possible. After using this tool, the author reviewed and edited the content as needed and takes full responsibility for the content of the publication.

Author Contributions: Conceptualization, W.L.; methodology, W.L.; software, W.L.; validation, W.L.; formal analysis, W.L.; investigation, W.L.; resources, W.L.; data duration, W.L.; writing–original draft preparation, W.L.; writing–review and editing, W.L.; visualization, W.L.; supervision, W.L.; project administration, W.L.; funding acquisition, not applicable.

Funding: This research received no external funding.

Conflicts of Interest: The author declares no conflict of interest.

 doi:10.20944/preprints202311.1048.v1

References

1. Fuji, H.; Qi, F.; Qu, L.; Takaesu, Y.; Hoshino, T. Prediction of Ligand Binding Affinity to Target Proteins by Molecular Mechanics Theoretical Calculation. *Chemical and Pharmaceutical Bulletin* **2017**, *65*, 461–468.
2. Berman, H.; Henrick, K.; Nakamura, H. Announcing the worldwide Protein Data Bank. *Nature Structural & Molecular Biology* **2003**, *10*, 980–980.
3. Li, W. Half-a-century Burial of ρ, θ and φ in PDB **2021**.
4. Li, W. How do SMA-linked mutations of *SMN1* lead to structural/functional deficiency of the SMA protein? *PLOS ONE* **2017**, *12*, e0178519.
5. Nishimura, E.; Pridal, L.; Glendorf, T.; Hansen, B.F.; Hubálek, F.; Kjeldsen, T.; Kristensen, N.R.; Lützen, A.; Lyby, K.; Madsen, P.; Pedersen, T.Å.; Ribel-Madsen, R.; Stidsen, C.E.; Haahr, H. Molecular and pharmacological characterization of insulin icodec: a new basal insulin analog designed for once-weekly dosing. *BMJ Open Diabetes Research & Care* **2021**, *9*, e002301.
6. Singh, A.K.; Singh, A.; Singh, R.; Misra, A. Once-weekly basal insulin icodec: Looking ONWARDS from pharmacology to clinical trials. *Diabetes & Metabolic Syndrome: Clinical Research & Reviews* **2022**, *16*, 102615.
7. Bajaj, H.S.; Bergenstal, R.M.; Christoffersen, A.; Davies, M.J.; Gowda, A.; Isendahl, J.; Lingvay, I.; Senior, P.A.; Silver, R.J.; Trevisan, R.; Rosenstock, J. Switching to Once-Weekly Insulin Icodec Versus Once-Daily Insulin Glargine U100 in Type 2 Diabetes Inadequately Controlled on Daily Basal Insulin: A Phase 2 Randomized Controlled Trial. *Diabetes Care* **2021**, *44*, 1586–1594.
8. Rosenstock, J.; Bain, S.C.; Gowda, A.; Jódar, E.; Liang, B.; Lingvay, I.; Nishida, T.; Trevisan, R.; Mosenzon, O. Weekly Icodec versus Daily Glargine U100 in Type 2 Diabetes without Previous Insulin. *New England Journal of Medicine* **2023**, *389*, 297–308.
9. Philis-Tsimikas, A.; Bajaj, H.S.; Begtrup, K.; Cailleteau, R.; Gowda, A.; Lingvay, I.; Mathieu, C.; Russell-Jones, D.; Rosenstock, J. Rationale and design of the phase 3a development programme (ONWARDS 1–6 trials) investigating once-weekly insulin icodec in diabetes. *Diabetes, Obesity and Metabolism* **2022**, *25*, 331–341.
10. Philis-Tsimikas, A.; Asong, M.; Franek, E.; Jia, T.; Rosenstock, J.; Stachlewska, K.; Watada, H.; Kellerer, M. Switching to once-weekly insulin icodec versus once-daily insulin degludec in individuals with basal insulin-treated type 2 diabetes (ONWARDS 2): a phase 3a, randomised, open label, multicentre, treat-to-target trial. *The Lancet Diabetes & Endocrinology* **2023**, *11*, 414–425.
11. Mathieu, C.; Ásbjörnsdóttir, B.; Bajaj, H.S.; Lane, W.; Matos, A.L.S.A.; Murthy, S.; Stachlewska, K.; Rosenstock, J. Switching to once-weekly insulin icodec versus once-daily insulin glargine U100 in individuals with basal-bolus insulin-treated type 2 diabetes (ONWARDS 4): a phase 3a, randomised, open-label, multicentre, treat-to-target, non-inferiority trial. *The Lancet* **2023**, *401*, 1929–1940.
12. Lingvay, I.; Asong, M.; Desouza, C.; Gourdy, P.; Kar, S.; Vianna, A.; Vilsbøll, T.; Vinther, S.; Mu, Y. Once-Weekly Insulin Icodec vs Once-Daily Insulin Degludec in Adults With Insulin-Naive Type 2 Diabetes. *JAMA* **2023**, *330*, 228.
13. Lingvay, I.; Buse, J.B.; Franek, E.; Hansen, M.V.; Koefoed, M.M.; Mathieu, C.; Pettus, J.; Stachlewska, K.; Rosenstock, J. A Randomized, Open-Label Comparison of Once-Weekly Insulin Icodec Titration Strategies Versus Once-Daily Insulin Glargine U100. *Diabetes Care* **2021**, *44*, 1595–1603.
14. Kjeldsen, T.B.; Hubálek, F.; Hjørringgaard, C.U.; Tagmose, T.M.; Nishimura, E.; Stidsen, C.E.; Porsgaard, T.; Fledelius, C.; Refsgaard, H.H.F.; Gram-Nielsen, S.; Naver, H.; Pridal, L.; Hoeg-Jensen, T.; Jeppesen, C.B.; Manfè, V.; Ludvigsen, S.; Lautrup-Larsen, I.; Madsen, P. Molecular Engineering of Insulin Icodec, the First Acylated Insulin Analog for Once-Weekly Administration in Humans. *Journal of Medicinal Chemistry* **2021**, *64*, 8942–8950.
15. Rosenstock, J.; Bajaj, H.S.; Janež, A.; Silver, R.; Begtrup, K.; Hansen, M.V.; Jia, T.; Goldenberg, R. Once-Weekly Insulin for Type 2 Diabetes without Previous Insulin Treatment. *New England Journal of Medicine* **2020**, *383*, 2107–2116.
16. Bajaj, H.S.; Goldenberg, R.M.; and. Insulin Icodec Weekly: A Basal Insulin Analogue for Type 2 Diabetes. *European Endocrinology* **2023**, *19*, 4.
17. DiMarchi, R.D.; Mayer, J.P. Icodec Advances the Prospect of Once-Weekly Insulin Injection. *Journal of Medicinal Chemistry* **2021**, *64*, 8939–8941.
18. Rosenstock, J.; Prato, S.D. Basal weekly insulins: the way of the future! *Metabolism* **2022**, *126*, 154924.

19. Kalra, S.; Bhattacharya, S.; Kapoor, N. Contemporary Classification of Glucagon-Like Peptide 1 Receptor Agonists (GLP1RAs). *Diabetes Therapy* **2021**, *12*, 2133–2147.
20. Pratley, R.; Amod, A.; Hoff, S.T.; Kadowaki, T.; Lingvay, I.; Nauck, M.; Pedersen, K.B.; Saugstrup, T.; Meier, J.J. Oral semaglutide versus subcutaneous liraglutide and placebo in type 2 diabetes (PIONEER 4): a randomised, double-blind, phase 3a trial. *The Lancet* **2019**, *394*, 39–50.
21. Anderson, S.L.; Beutel, T.R.; Trujillo, J.M. Oral semaglutide in type 2 diabetes. *Journal of Diabetes and its Complications* **2020**, *34*, 107520.
22. Li, W. Strengthening Semaglutide-GLP-1R Binding Affinity via a Val27-Arg28 Exchange in the Peptide Backbone of Semaglutide: A Computational Structural Approach. *Journal of Computational Biophysics and Chemistry* **2021**, *20*, 495–499.
23. Nadkarni, P.; Chepurny, O.G.; Holz, G.G. Regulation of Glucose Homeostasis by GLP-1. In *Progress in Molecular Biology and Translational Science*; Elsevier, 2014; pp. 23–65.
24. Lau, J.; Bloch, P.; Schäffer, L.; Pettersson, I.; Spetzler, J.; Kofoed, J.; Madsen, K.; Knudsen, L.B.; McGuire, J.; Steensgaard, D.B.; Strauss, H.M.; Gram, D.X.; Knudsen, S.M.; Nielsen, F.S.; Thygesen, P.; Reedtz-Runge, S.; Kruse, T. Discovery of the Once-Weekly Glucagon-Like Peptide-1 (GLP-1) Analogue Semaglutide. *Journal of Medicinal Chemistry* **2015**, *58*, 7370–7380.
25. Gabery, S.; Salinas, C.G.; Paulsen, S.J.; Ahnfelt-Rønne, J.; Alanentalo, T.; Baquero, A.F.; Buckley, S.T.; Farkas, E.; Fekete, C.; Frederiksen, K.S.; Helms, H.C.C.; Jeppesen, J.F.; John, L.M.; Pyke, C.; Nøhr, J.; Lu, T.T.; Polex-Wolf, J.; Prevot, V.; Raun, K.; Simonsen, L.; Sun, G.; Szilvásy-Szabó, A.; Willenbrock, H.; Secher, A.; Knudsen, L.B. Semaglutide lowers body weight in rodents via distributed neural pathways. *JCI Insight* **2020**, *5*.
26. Plum-Mörschel, L.; Andersen, L.R.; Hansen, S.; Hövelmann, U.; Krawietz, P.; Kristensen, N.R.; Lehrskov, L.L.; Haahr, H. Pharmacokinetic and Pharmacodynamic Characteristics of Insulin Icodec After Subcutaneous Administration in the Thigh, Abdomen or Upper Arm in Individuals with Type 2 Diabetes Mellitus. *Clinical Drug Investigation* **2023**, *43*, 119–127.
27. Belal, H.; Gandhi, G.Y. In uncontrolled T2DM treated with a basal-bolus insulin regimen, weekly icodec was noninferior to daily glargine for HbA1c at 26 wk. *Annals of Internal Medicine* **2023**, *176*, JC94.
28. Anderson, S.L.; Bassetti, M.; Mangoni, A.A. Drugs in Context Editorial: Review of 2020 and what lies ahead in therapeutic interventions. *Drugs in Context* **2021**, *10*, 1–5.
29. Griffin, T.P.; Dinneen, S.F. In T2DM, weekly insulin icodec did not differ from daily glargine for reducing HbA1c or significant/severe hypoglycemia. *Annals of Internal Medicine* **2021**, *174*, JC34.
30. Pieber, T.R.; Arfelt, K.N.; Cailleteau, R.; Hart, M.; Kar, S.; Mursic, I.; Svehlikova, E.; Urschitz, M.; Haahr, H. Hypoglycaemia frequency and physiological response after double or triple doses of once-weekly insulin icodec vs once-daily insulin glargine U100 in type 2 diabetes: a randomised crossover trial. *Diabetologia* **2023**, *66*, 1413–1430.
31. Zerihun, K.; Mhanna, M.; Ayesh, H.; Ghazaleh, S.; Khader, Y.; Beran, A.; Aldhafeeri, A.; Sharma, S.; Iqbal, A.; Legesse, H.; Jaume, J. Efficacy and Safety of Insulin Icodec Versus Glargine U100: A Meta-Analysis of Randomized Controlled Trials. *American Journal of Therapeutics* **2022**, *30*, e480–e483.
32. Bellary, S.; Barnett, A.H. Insulin icodec: evolution or revolution in diabetes therapy? *The Lancet Diabetes & Endocrinology* **2023**, *11*, 379–380.
33. e Silva, R.R.; de Miranda Gauza, M.; Guisso, M.E.S.; da Silva, J.O.N.; Kohara, S.K. Once-Weekly Insulin Icodec vs. Once-Daily Insulin Glargine U100 for type 2 diabetes: a systematic review and meta-analysis of phase 2 randomized controlled trials. *Archives of Endocrinology and Metabolism* **2023**, *67*.
34. Pieber, T.R.; Asong, M.; Fluhr, G.; Höller, V.; Kristensen, N.R.; Larsen, J.H.; Ribel-Madsen, R.; Svehlikova, E.; Vinther, S.; Voortman, M.; Haahr, H. Pharmacokinetic and pharmacodynamic properties of once-weekly insulin icodec in individuals with type 2 diabetes. *Diabetes, Obesity and Metabolism* **2023**.
35. Russell-Jones, D.; Babazono, T.; Cailleteau, R.; Engberg, S.; Irace, C.; Kjaersgaard, M.I.S.; Mathieu, C.; Rosenstock, J.; Woo, V.; Klonoff, D.C. Once-weekly insulin icodec versus once-daily insulin degludec as part of a basal-bolus regimen in individuals with type 1 diabetes (ONWARDS 6): a phase 3a, randomised, open-label, treat-to-target trial. *The Lancet* **2023**.
36. Feher, J. Digestion and Absorption of the Macronutrients. In *Quantitative Human Physiology*; Elsevier, 2017; pp. 821–833. doi:10.1016/b978-0-12-800883-6.00081-1.

37. Gutmann, T.; Schäfer, I.B.; Poojari, C.; Brankatschk, B.; Vattulainen, I.; Strauss, M.; Ünal Coskun. Cryo-EM structure of the complete and ligand-saturated insulin receptor ectodomain. *Journal of Cell Biology* **2019**, *219*.
38. Waterhouse, A.; Bertoni, M.; Bienert, S.; Studer, G.; Tauriello, G.; Gumienny, R.; Heer, F.T.; de Beer, T.A.P.; Rempfer, C.; Bordoli, L.; Lepore, R.; Schwede, T. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Research* **2018**, *46*, W296–W303.
39. Biasini, M.; Bienert, S.; Waterhouse, A.; Arnold, K.; Studer, G.; Schmidt, T.; Kiefer, F.; Cassarino, T.G.; Bertoni, M.; Bordoli, L.; Schwede, T. SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Research* **2014**, *42*, W252–W258.
40. Webb, B.; Sali, A. Protein Structure Modeling with MODELLER. In *Methods in Molecular Biology*; Springer US, 2020; pp. 239–255.
41. DeLano, W.L. Pymol: An open-source molecular graphics tool. *CCP4 Newsletter On Protein Crystallography* **2002**, *40*, 82–92.