Designing novel oral-Insulin conjugates for the development of oral-Insulin tablet: Inulin-Insulin conjugate is an efficient form for oral-Insulin tablet

Publication History
Received: 28 July 2015
Accepted: 07 September 2015
Published: 1 October 2015

Citation
Seenivasagam R. Designing novel oral-Insulin conjugates for the development of oral-Insulin tablet: Inulin-Insulin conjugate is an efficient form for oral-Insulin tablet. Drug Discovery, 2015, 10(26), 121-249
Designing novel oral-Insulin conjugates for the development of oral-Insulin tablet: Inulin-Insulin conjugate is an efficient form for oral-Insulin tablet

Seenivasagam R1,*

1. Department of Biomedical Science, India

*Corresponding Author: Seenivasagam R; Department of Biomedical Science, India; Email: Seenivasagam.pharma@gmail.com

ABSTRACT

Insulin is used medically to manage diabetes mellitus by frequent subcutaneous injections have turned into a part of life can be enormously distressing for patients. Hence, we anticipate to getting rid of the usage of subcutaneous injections completely. Virtual screening method (for selection of carriers), Toxtree tool (for toxicity evaluation of carriers) and Discovery Studio tool (for Pharmacopore designing, ADMET analysis, designing oral Insulin conjugates and Interaction studies between Insulin Receptor and oral Insulin conjugates) were used for proposed study. We have screened 14 competent drugs delivering agents (DDAs) from 7 chemical compound databases. The ADMET and Pharmacoporic properties of DDAs were analyzed by drug-informatics' tools. Consequently, the DDAs were mono, di & poly conjugated by covalent bonding with various binding sites of Monomeric and hexameric form of human insulin and insulin-lispro (Humalog®) individually; and novel oral-insulin conjugates (OICs) were generated. Its binding efficiency and biological activity with Insulin-receptor were determined. Inulin and Vitamin-B1 are considered as novel, safe and proficient carriers for oral delivery of Insulin. Insulin Lispro is the remarkable option for oral delivery than other Insulin forms.

Keywords: Oral Insulin: Diabetes; Drug delivery; Insulin tablets

Abbreviations: DDAs – Drug delivering agents; OICs – Oral Insulin Conjugates; Da - Dalton; FDA – Food & Drug Administration; IN-105 - Methoxy-poly(ethylene glycol)-insulin conjugates; HIM2 - Hexyl-insulin monocunjugate-2; KEGG - Kyoto Encyclopedia of Genes and Genomes; ChEBI - Chemical Entities of Biological Interest; log(Sw) - log Aqueous solubility (Solubility in water); logBB - log Brain-Blood; CYP2D6 - cytochrome P450 2D6; HIA - Human intestinal absorption; PSA - polar surface area; TOPKAT - Toxicity Prediction by Komputer Assisted Technology; BFGS - Broyden-Fletcher-Goldfarb-Shanno; BBB- Blood brain barrier; ADMET - Absorption, Distribution, Metabolism, Excretion and Toxicity; SASA - Solvent Accessible Solvent Area; MTD - Maximum tolerated dose; C1 - Vitamin B12 (cobalamin); C2 - Vitamin H (Biotin); C3 - Folic acid (Vitamin M / Vitamin B9); C4 - Vitamin B1 (Thiamin); C5 - L-Carnitine (Vitamin BT); C6 - Poly-N-vinylpyrrolidone; C7; Inulin; C8 - Poly Cysteine; C9 - Chitosan; C10 - Pectin; C11 - Poly(propylene glycol); C12 - Poly(propylene imine); C13 - Poly (lactic-co-glycolic acid); C14 - Deoxycholic acid.

1. INTRODUCTION

Drug delivery is an essential process which leads to innovation of novel carriers to manage oral delivery of some peculiar drugs such as proteins, hormones, enzymes, and peptides. These insulin drugs are requires efficient drug delivering agents (DDAs) for oral formulations to reach therapeutic targets against the interruption of biological barriers such as gastric digestion in stomach, proteolysis in intestinal regions, and poor bioavailability due to higher molecular weight. Now a day, the subcutaneous injection of insulin formulations has been the core therapy utilized for the management of type-1 diabetes as well as type-2 diabetes since insulin’s discovery over 85 years ago. Multiple doses of insulin subcutaneous injections on the daily basis making patients painful and create ‘fear of pain’ among them. There are several disadvantages of this subcutaneous administration such as poor compliance, local discomfort, inconvenience, fright of soreness of multiple injections, hypoglycemic risk related to injections, occasional hyper-insulinemia due to overdoses, unsatisfactory metabolic regulations, allergy, harrowing, and insulin lipodystrophy at the site of injection. In addition, subcutaneously administered insulin is absorbed directly into the peripheral circulation without initial hepatic circulation, thereby exposing peripheral targets to higher insulin concentrations relative to the liver. As of 2014, no products appeared to be successful in the market because of lack of physical stability against biological barriers. For those effective reasons, we intend to screen novel DDAs to design efficient oral Insulin conjugates (OICs) against diabetes.
In our analysis, the monomeric & hexameric form of recombinant human insulin and Humulin were modified by the covalent bonding of short-chain DDAs. The modification takes place to the either or both amino-acids of PheB1 & LysB29 in recombinant human insulin, and also PheB1 & LysB28 in Humulin. Because, the conjugation of DDAs to Insulin forms ought to have no negative impact on the insulin’s therapeutic activity since it has been formerly determined that these residues do not directly contribute in receptor binding. 10-14 On these perspectives, a study proven that DDA-recombinant Insulin conjugate (at LysB29) was physically more stable against acidic condition, not undergoes enzymatic degradation, less immunogenic, less antigenic, resistance to fibrillation in aqueous solution and improved stability against temperature, pH, and interfacial & shear forces15-18 when correlated with DDA-recombinant Insulin conjugate (at PheB1), and human native-Insulin.8,9,19 Moreover, the conjugation of recombinant insulin with DDA can protect the self-aggregation (dimerization) in aqueous solutions.15 In this stance, Humulin has higher resistant capability against self-aggregation & other biological barriers than human recombinant-insulin and native-insulin.20 The oligomeric-insulin forms (chiefly hexamers) is not bioactive and the fraction of an amount is absorbed across the capillary endothelium into the systemic circulation in the absence of DDA. The dissociation of oligomer into dimmers and monomers is seen as the rate-limiting barrier to absorption that effectively affects the preparation’s pharmacological response.21 Meantime, the monomeric-Insulin forms of is highly bio-available but easily de-nature & de-folded in presence of biological barriers.15 In a study, the in-vivo pharmacodynamic assay reveals that there is no loss of biological activity after conjugation of carrier to the either site on the oligomeric form of insulin-B-chain.21 On the other hand, the attachment of long-chain DDA (2000Da) decreased the bioactivity of conjugates than sort-chain DDA (750Da). Hence, mono-disperse and short-chain DDAs (≤750Da) are preferred.21 From the earlier studies, we choose that Humulin, monomeric & hexameric form of human recombinant insulin and short chain DDAs for designing OICs. We report, and analyze the structural series of novel DDAs and OICs by drug-informatics.

2. Methods
2.1. Carriers
2.1.1. Screening
For Virtual Screening22 of DDAs, PubChem Compound,23 Zinc Database,24 KEGG25 ([Kyoto Encyclopedia of Genes and Genomes]), DrugBank,26 ChemSpider,27 ChEMBL,28 and ChEBI29 (Chemical Entities of Biological Interest), compound databases were used. Carrier agents were retrieved by the search terms of "Polymer" & "Biopolymer". The compounds were filtered through virtual screening & biomedical text mining using a defined criteria listed below, in part akin to the Lipinski’s rule:30 the DDAs should be : a) mono-disperse, b) short-chain (low-molecular weight (<2000Da)), c) biocompatible, d) lipophilic, e) physically stability against gastric acids and proteolytic enzymes, f) inert (no biological activity), g) non-toxic.

2.1.2. Analysis of Pharmacophoric features
The retrieved compounds were subjected to Pharmacophore analysis by Discovery Studio (Accelrys discovery studio 2.5).31 In Discovery Studio, a pharmacophore is defined as the essential features or chemical substructures and their corresponding 3D locations that are responsible for the similar biological activities of a set of compounds. Typically, pharmacophore features include hydrophobic (in light blue), hydrogen bond acceptor (HBA, in green), hydrogen bond donor (HBD, in Magenta), and active principles.

2.1.3. Analysis of Physico-chemical properties
The retrieved compounds were subjected to ADMET evaluation by Discovery Studio. ADMET Descriptors include:

- **Aqueous Solubility**: This model uses linear regression to predict the solubility of each compound in water at 25°C. Key to aqueous solubility is graded through Level, Value and Drug-likeliness as follows: (level 0; log(Sw) < -8.0; Extremely low), (level 1; -8.0 < log(Sw) < -6.0; No, very low, but possible), (level 2; -6.0 < log(Sw) < -4.0; Yes, low), (level 3; -4.0 < log(Sw) < -2.0; Yes, good), (level 4; -2.0 < log(Sw) 0.0=™; Yes, optimal), (level 5; 0.0 < log(Sw); No, too soluble).
- **Blood Brain Barrier Penetration**: This model predicts blood-brain penetration (blood brain barrier, BBB) after oral administration. This model contains a quantitative linear regression model for the prediction of blood-brain penetration, as well as 95% and 99% confidence ellipses in the ADMET_PSA_2D, ADMET_AlogP98 plane (ADMET_PSA_2D means Fast polar surface area; ADMET_AlogP98 means Atom-based LogP). There are four prediction levels within the 95% and 99% confidence ellipsoids and they are graded through Level, Value and Brain-Blood ratio as follows: (Level 0; Very high penetrants (logBB ≥ 0.7); Brain-Blood ratio greater than 5:1), (Level 1; High penetrants (0 ≤ logBB < 0.7); Brain-Blood ratio between 1:1 and 5:1), (Level 2; Medium penetrants (-0.52 < logBB < 0); Brain-Blood ratio between 0.3:1 and 1:1), (Level 3; Low penetrants (logBB ≤ -0.52); Brain-Blood ratio less than 0.3:1), (Level 4; Undefined; Outside 99% confidence ellipse).
- **CYP2D6 Binding**: Predicts cytochrome P450 2D6 enzyme inhibition. The cytochrome P450 2D6 model predicts CYP2D6 enzyme inhibition using 2D chemical structure as input. The model classifies compounds as either 0 or 1 for non-inhibitor or inhibitor and provides an average-class-value estimate of confidence. Key to CYP2D6 is graded through Predicted class, value and Description as follows: (Predicted class 0; Non-inhibitor; Unlikely to inhibit CYP2D6 enzyme; ADMET_CYP2D6_Probability < 0.5), (Predicted class 1; Inhibitor; Likely to inhibit CYP2D6 enzyme; ADMET_CYP2D6_Probability > 0.5). ADMET_CYP2D6_Probability means CYP2D6 score or average class value.
- **Hepatotoxicity**: Predicts the occurrence of dose-dependent human hepatotoxicity. The hepatotoxicity model predicts potential organ toxicity for a wide range of structurally diverse compounds. Key to Hepatotoxicity is graded through Predicted class, value and
Description as follows; (Predicted class 0; Nontoxic; Unlikely to cause dose-dependent liver injuries. ADMET_Hepatotoxicity_Probability < 0.5), (Predicted class 1; Toxic; Likely to cause dose-dependent liver injuries. ADMET_Hepatotoxicity_Probability > 0.5). ADMET_Hepatotoxicity_Probability means Hepatotoxicity score (average-class value).

- Intestinal Absorption: This model predicts human intestinal absorption (HIA) after oral administration. Intestinal absorption is defined as a percentage absorbed rather than as a ratio of concentrations (cf. blood-brain penetration). A well-absorbed compound is one that is absorbed at least 90% into the bloodstream in humans. The intestinal absorption model includes 95% and 99% confidence ellipses in the ADMET_PSA_2D, ADMET_AlogP98 plane. The ellipses define regions where well-absorbed compounds are expected to be found: 95% of well-absorbed compounds are expected to fall within the 95% ellipse, while 99% of well-absorbed compounds should fall within the 99% ellipse. Note that the location of any particular compound does not necessarily imply whether it will be well, moderately or poorly absorbed. In general, however, absorption tends to drop off quite rapidly outside the 95% ellipse. These levels are defined by the 95% (blue line) and 99% (magenta line) confidence ellipsoids. There are four prediction levels and they are graded through level, value and Description as follows: (Level 0; ADMET_Absorption_T2_2D < 6.1261 (inside 95%); Good absorption), (Level 1; 6.1261 ≤ ADMET_Absorption_T2_2D < 9.6026 (inside 99%); Moderate absorption), (Level 2; 9.6026 < ADMET_Absorption_T2_2D (outside 99%); Low absorption), (Level 3; ADMET_PSA_2D ≥ 150.0 or ADMET_AlogP98 ≤ -2.0 or ADMET_AlogP98 ≥ 7.0; Very low absorption). ADMET_Absorption_T2_2D is the Mahalanobis distance for the compound in the ADMET_PSA_2D, ADMET_AlogP98 plane. It is referenced from the center of the region of chemical space defined by well-absorbed compounds.

- Plasma Protein Binding: The plasma protein binding model predicts whether a compound is likely to be highly bound to carrier proteins in the blood. Key to Plasma Protein Binding is graded through level and Description as follows; (Level 0; Binding is < 90% (No markers flagged and AlogP98 < 4.0)), (Level 1; Binding is > 90% (flagged at 90% or AlogP98 > 4.0)), (Level 2; Binding is > 95% (flagged at 95% or AlogP98 > 5.0)). AlogP98 means Atom-based LogP from FastDesc.

2.1.4. Analysis of Toxicity

The retrieved compounds were subjected to Toxicity evaluation by Discovery Studio and TOXTREE (by IdeaConsult Ltd (Sofia, Bulgaria)). Toxtree is able to estimate toxic hazard by applying a decision tree approach. The classification result is shown in graphical form (green highlight for class I (non-toxic), yellow highlight for class II (Moderately toxic) and red highlight for class III (Toxic)), as well as in text form. In Discovery Studio; TOPKAT models have been re-trained using updated training sets from the legacy TOPKAT (Toxicity Prediction by Komputer Assisted Technology). The following models are extensible and are derived using calculable properties:

- FDA Rodent Carcinogenicity
- Ames Mutagenicity
- Rat Oral LD50
- Rat Maximum Tolerated Dose
- Skin Irritancy
- Skin Sensitization
- Aerobic Biodegradability

2.2. Designing Oral insulin conjugates & OIC – IR binding

For designing, Oral insulin conjugates & Interaction of OIC with IR were carried out through “LibDock” algorithm of Discovery Studio. The LibDock docking program performs the following steps using a set of pre-generated ligand conformations and a receptor with a specified binding site:

- Remove hydrogen atoms.
- Rank ligand conformations and prune by Solvent Accessible Solvent Area (SASA).
- Find hotspots using a grid placed into the binding site and using polar and apolar probes. The numbers of hotspots are pruned by clustering to a user defined value.
- Dock ligand poses by aligning to the hotspots. This is performed by using triplets (i.e., three ligand atoms are aligned to three receptor hotspots). Poses which result in protein clashes are removed.
- A final Broyden-Fletcher-Goldfarb-Shanno (BFGS) pose optimization stage is performed using a simple pair-wise score (similar to Piecewise Linear Potential). The top scoring ligand poses are retained.
- Hydrogen atoms are added.

Hydrogen atoms added in the final step may result in small bumps with the protein. Therefore, minimization should be performed prior to using scoring functions that are sensitive to such bumps.

3. RESULTS & DISCUSSION

3.1. Carriers

3.1.1. Screening

In carrier screening (Flowchart 1), more than 1, 00,000 compounds were retrieved from 7 compound databases. Among those, 14 compounds were screened by using experimental text mining & filtration criteria (Table 1). According to data mining, most of the screened compounds in Table 1 are monodisperse such as Vitamin B12, Vitamin H, Folic acid, Poly-N-vinylpyrrolidone, Inulin, Poly Cysteine, Chitosan.
Pectin, Poly (Propylene glycol), Poly (Propylene imine), Poly (lactic-co-glycolic acid), Deoxycholic acid except Vitamin B1 and L-Carnitine, because of the lack of experimental data. Molecular weight of polymeric drug delivering molecules are varies based on length of chain, but in the case of Vitamins, molecular weights are measurable. Maximum carriers in the retrieved list have shown low-molecular weight due to short-chain in structure (<2000 Daltons). L-Carnitine (162.113 Daltons), Poly-N-vinylpyrrolidone (11.141 Daltons), Inulin (342.297 Daltons), Poly Cysteine (121.158 Daltons), Pectin (194.139 Daltons) and Poly(propylene glycol) (76.094 Daltons) are possess low-molecular weight while compare with vitamins and other macromolecules in the list. All the listed molecules are biocompatible and biodegradable.

Most of the carriers are having efficient oral bioavailability and intestinal permeability such as Vitamin B12, Folic acid, Vitamin H, Poly Cysteine, Poly (lactic-co-glycolic acid), and Deoxycholic acid. Vitamin B1, L-Carnitine, Inulin, Poly(propylene glycol), Poly(propylene imine), Poly (lactic-co-glycolic acid) and Deoxycholic acid are graded as moderately efficient in bioavailability while Poly-N-vinylpyrrolidone and Poly Cysteine are very poor intestinal transport, because of the lack of lipophilicity. Biomedical text mining shows all the listed compounds are physically stable against gastric acids and proteolytic enzymes during drug delivery. Among the retrieved carriers, Vitamin B12, Vitamin H, Vitamin B1, Poly-N-vinylpyrrolidone, Inulin, Poly(propylene glycol), Poly (lactic-co-glycolic acid) and Deoxycholic acid are chemically inert and they do not undergoes any biochemical transformation & aggregation during drug delivery. Through text-mining, we could not found whether Folic acid, L-Carnitine, Poly Cysteine, Pectin, and Poly(propylene imine) are inert or not. Based on the concept of toxicity, Vitamin B12, Vitamin H, Folic acid, Vitamin B1, L-Carnitine, Inulin, Poly Cysteine, Chitosan, Pectin, Poly (lactic-co-glycolic acid) and Deoxycholic acid are non-toxic materials and does not produce any untoward reactivity during drug-delivery mechanism. But, Poly-N-vinylpyrrolidone, Poly(propylene glycol), Poly(propylene imine), and Deoxycholic acid are moderately toxic based on text-mining.

The experimental text-mining concludes Inulin, Chitosan and Poly (lactic-co-glycolic acid) are efficient, safe and primary drug delivering molecules of drugs that completely fulfill the filtration criteria. Inulin reduces the production of potentially toxic metabolites, induce important immune-mediated effects and reduce the cancer risk during drug delivery. Generally Inulin and Chitosan are participating in colon-targeting drug delivery, possess very minimum exposure to gastric fluids in the stomach and enzymatic degradation in the small intestine. Folic acid, L-Carnitine and Pectin are eligible for drug delivery process, but we could not conclude whether they are inert. Vitamin B1 is an effectual and non-toxic carrier, may cause Anaphylaxis at higher doses. L-Carnitine & Vitamin B1 does not hold any literature evidence that they possess mono-disperse character or not. Vitamin B12 is inert, long-term usage may cause Vit-B12 deficiency. Vitamin H is inert, but it may impart the biological activity of the drug. Poly-N-vinylpyrrolidone, and Poly (propylene imine) are participating in nanoparticle-based drug delivery, and both are moderately toxic. Poly Cysteine is hydrophilic in nature; shows poor permeability across intestinal epithelium; and no research confirmation whether it is inert. Poly(propylene glycol) is an efficient carrier for drug delivery, and inert, overdose may cause skin & eye irritation. It may induce Cardiotoxic effects include arrhythmias and cardiac arrest, CNS depression, Renal and hepatic damage has also reported. In case studies, toxic symptoms appeared only after frequent doses of propylene glycol, used as a vehicle for medicines, were repetitively applied to the skin. In drug delivery mechanism, it eliminates toxic degradation. It is less toxic than the parent substance (Poly ethylene glycol). Deoxycholic acid is a prominent carrier, which is not inert and moderately toxic; may promote colon tumorigenesis in both animals and humans.

3.1.2. Analysis of Pharmacophoric features
Pharmacophores are conceptual description of molecular principles or features that are essential for molecular recognition of a molecule through a biological macromolecule. Pharmacophoric features of Drug delivering molecules are illustrated and demonstrated by Discovery Studio software (Table 2: Figure 1). Pharmacophore features comprise hydrogen bond donor, hydrogen bond acceptors, hydrophobic centroids, aromatic rings, cations, and anions. Among those features, Acceptor, Donor, hydrophobic regions and the number of active principles of carrier molecules (Supplementary Figure 1) were investigated based on Lipinski’s rule. According to the rule of Lipinski, a molecule should not donate more than 5 hydrogen bonds and it should not accept more than 10 hydrogen bonds. Most of the carriers obey the rule and eligible for drug delivery except Inulin and Vitamin B12. Inulin posses 11 hydrogen bond acceptors (≤ 10 as per rule) may leads to accepting electrons transferred to it from another compound. Inulin may be an oxidizing molecule, by virtue of its accepting electrons, it itself reduced in the drug delivery process. Inulin may undergo permanent chemical alteration through covalent or ionic reaction chemistry, resulting in the complete and irreversible transfer of one or more electrons. But in this case, Inulin is a colon targeting molecule; subsequently it does not undergo any chemical transformation during drug delivery mechanism like Chitosan. Vitamin B12 posses 7 hydrogen bond donors (≤5 as per rule) may leads to donate electrons to another compound. Vitamin B12 may be a reducing agent, by virtue of its donating electrons, it itself oxidized in the drug delivery process. Compare with another Drug delivering molecules, Inulin, Vitamin B12 and Deoxycholic acid contains maximum hydrophobic regions (Table 2) that is responsible for efficient intestinal transport. The significant criteria, pharmacophoric active principle is more in Folic acid, Vitamin B1, Inulin, Chitosan and Pectin, But Vitamin B12 and Poly(propylene imine) does not hold any single active principle in their structure. The analysis of pharmacophoric features suggests that Vitamin B1, Inulin, Chitosan, and Pectin are possible Drug delivering molecules for Insulin.

3.1.3. Analysis of Physico-chemical properties
ADME descriptors of Drug delivering molecules were evaluated by Discovery Studio and listed in Table 3 (Supplementary Figure 2). Vitamin H, Vitamin B1, Poly Cysteine, Poly(propylene glycol), Poly (lactic-co-glycolic acid) and Deoxycholic acid are falls within 95% absorption ellipse that shows efficient absorption (level = 0) across intestinal epithelium, concurrently shows the sign of maximum bioavailability in drug delivering
mechanism. Poly-N-vinylpyrrolidone and Poly(propylene imine) are falls within 99% absorption ellipse that shows moderate absorption (level = 1); remaining compounds are falls outside 99% absorption ellipse shows poor absorption (level = 3). In the case of Aqueous solubility, Folic acid (-3.378) show extremely high Aqueous Solubility and drug-likeness. Vitamin H (-1.432), Vitamin B1 (-1.335), Poly-N-vinylpyrrolidone (-0.550) and Poly(propylene imine) (-0.097) shows optimal Aqueous Solubility and drug-likeness; others show poor drug-likeness. In the point of Blood-Brain-barrier (BBB) penetration, Deoxycholic acid (-0.154) is medium penetrant across Blood-Brain-barrier, because it is fall within 99% confidence ellipsoids (level = 2), the Brain-Blood ratio is between 0.3:1 and 1:1. Vitamin H (-1.229), Vitamin B1 (-1.253), Poly Cysteine (-1.362), Poly(propylene glycol) (-1.039), Poly (lactic-co-glycolic acid) (-1.713) are low penetrants across Blood-Brain-barrier, because it is within 99% confidence ellipsoids (level = 3), the Brain-Blood ratio is less than 0.3:1. Other carriers are poor penetrants because they are outside the 95% and 99% confidence ellipsoids (undefined level = 4). In the case of CYP-4502D6 binding, the carrier should be non-inhibitor because CYP-4502D6 is responsible for metabolism and elimination of drug molecules. All the listed drug delivering molecules are non-inhibitors because their ADMET CYP2D6 Probability is < 0.5. In the case of Hepatotoxicity, most of the selected carriers are non-toxic because, their ADMET hepatotoxicity probability is < 0.5. But Vitamin B12 (0.509) and Folic acid (0.662) are toxic because, their ADMET hepatotoxicity probability is > 0.5. Plasma Protein Binding capability should be <90%, then only the unbound molecule can easily penetrate the tissues to reach the active site and then to get eliminate. The Plasma Protein Binding character of all listed carriers is low and satisfies the standard value (< 4.0). Based on the overall results of Physico-chemical properties, Vitamin H, Vitamin B1, and Deoxycholic acid are superior drug delivering molecules according to their efficient solubility in liquid, moderate penetration across Blood Brain barrier, enhanced absorption across intestinal epithelial cells, non-inhibition of CYP-4502D6 binding, unbound nature with plasma proteins and low hepatotoxicity.

3.1. Analysis of Toxicity
Toxic scale is the most significant prediction for carriers in drug delivery mechanism. Toxicity properties such as FDA Rodent Carcinogenicity, Mutagenicity, Rat oral LD50 (kg Kg Body weight), Rat maximum tolerated dose (g/Kg Body weight), Skin Irritant, Skin sensitization, Aerobic biodegradability and general toxicity were studied by Discovery Studio and Toxtree (Table 4 & Supplementary Figure 3 & 4). The listed carriers are Non-mutagen and are Non-carcinogen except Vitamin B1 and Poly-N-vinylpyrrolidone. The predicted skin irritancy is severe for Deoxycholic acid, while Vitamin H, Folic acid, Vitamin B1, Poly Cysteine, Pectin, and Poly (lactic-co-glycolic acid) are non-irritants and others are in the category of mild-irritant. Vitamin B1 is predicted under strong skin-sensitizer, whereas Folic acid and Deoxycholic acid are weak skin-sensitizers; others fall under non-skin-sensitizer. The drug delivering molecules should be biodegradable in an aerobic environment after delivering the drug to receptors; otherwise it may leads to untoward reactivity. In this case, Folic acid, Vitamin B1 and Poly(propylene imine) are Non-Degradable, others are Degradable under aerobic condition. The TOPKAT model predicts the rat oral acute median lethal dose (LD50) in the toxicity test, and the rat maximum tolerated dose (MTD) of all drug delivering molecules. According to general toxicity prediction by Toxtree tool, Vitamin B12, Folic acid, Poly-N-vinylpyrrolidone, Chitosan, Poly(propylene glycol), Poly(propylene imine), and Deoxycholic acid may highly toxic while others are low in toxic category. Based on the analysis of toxicity studies, we concluded that Vitamin H, Inulin, Poly Cysteine, Pectin, and Poly (lactic-co-glycolic acid) are safe carriers for delivering oral insulin molecule.

3.2. Oral insulin conjugates
3.2.1. Designing
Human Insulin Monomer (PDB ID: 1HLS), human insulin hexamer (PDB ID: 1AIO), and Insulin Lispro (PDB ID: 1LPH) were retrieved from Protein data bank and conjugated with all carriers individually. Nikhil J Kavimandan et al. (2006) and Hinds et al. (2000) suggest that, conjugation of carriers with B1Phe, B27Thr, B28Pro & B29Lys amino acids of human Insulin and B28Lys amino acid of Insulin Lispro will be the efficient conjugate against ADMET barriers in oral delivery. Based on our computational analysis (Discovery Studio - LibDock), the positive LibDock score indicating the competent oral-insulin conjugation.

Human Insulin Monomer (PDB ID: 1HLS), conjugated individually with all listed drug delivering molecules (Table 5 & Supplementary Figure 5); among those Inulin was mono-conjugated efficiently with B1Phe of Insulin Monomer, and Poly (lactic-co-glycolic acid) was di-conjugated competitively with B1Phe & A11Cys of Insulin Monomer. Vitamin B12, Vitamin M, poly-N-vinylpyrrolidone, poly(propylene imine), and Deoxycholic acid did not shows any conjugation with amino acids of Insulin Monomer. Rest of the drug delivering molecules form the incompetent mono & di-conjugates with A & B chain amino acids of Insulin Monomer. Based on the Binding site (PHE B1) and LibDock score, Inulin (117.663) shows the competent oral-insulin conjugation (Figure 2c). Based on the LibDock score, Vitamin H (86.2835) and Vitamin B1 (94.1144) shows the competent oral-insulin conjugation (Figure 2a & 2b); but in the case of Poly (lactic-co-glycolic acid), the score is poor; even it conjugate at PHE B1 (Figure 2d).

Human insulin hexamer (PDB ID: 1AIO) conjugated individually with all listed drug delivering molecules (Table 6 & Supplementary Figure 6); among those Vitamin B1 and Inulin were mono-conjugated efficiently with B29Lys of Insulin hexamer and form the competent mono conjugates. Vitamin B12, poly-N-vinylpyrrolidone, and poly(propylene imine) did not shows any conjugation with amino acids of Insulin hexamer. Rest of the drug delivering molecules forms the incompetent mono & di-conjugates with A, B, C, D, E, F, G, H, I, J, K and L chain amino acids of Insulin hexamer. According to LibDock score, Vitamin M (114.324), Vitamin H (103.231) and Inulin (94.3543) show the competent oral-insulin conjugation (Figure 3a, 3b & 3c). Based on the Binding site (LYS B29), Inulin and Vitamin B1 shows the competent oral-insulin conjugation (Figure 3c & 3d).

Insulin Lispro (PDB ID: 1LPH) conjugated individually with all listed drug delivering molecules (Table 7 & Supplementary Figure 7); among those Vitamin M was poly conjugated with A & B chain amino acids of Insulin Lispro such as A1Gly, A2Ile, A3Val, A4Glu, A19Tyr, B4Gln, B5His, B27Thr, B28Lys, B30Thr. This conjugates efficiently bonding with B27Thr and B28Lys amino acids. It also made the inefficient multiple bonding
interactions with Insulin Lispro and may affect the absorption and bioavailability of Insulin Lispro. Vitamin B1 was di-conjugated reasonably with Lys B2 & Glu A4 of Insulin Lispro. Insulin was tetra-conjugated fairly with B2Lys, A1Gly, A3Val, and A4Glu of Insulin Lispro. Vitamin B12, Vitamin H, poly-N-vinylpyrrolidone, poly(propylene imine), and Deoxycholic acid did not shows any conjugation with amino acids of Insulin Lispro. Rest of the drug delivering molecules forms the incompetent mono, di, tetra & poly-conjugates with A & B chain amino acids of Insulin Lispro. According to LibDock score and Binding site (Lys B28), Vitamin M (131.57), Vitamin B1 (89.8971) and Inulin (76.2195) shows the competent oral-insulin conjugation (Figure 4a, 4b & 4c).

3.3. Interaction of Oral insulin conjugates with Insulin Receptor (IR)
Among the designed conjugates, Inulin-Insulin Monomer conjugate, Vitamin B1-Insulin Monomer conjugate, Vitamin H-Insulin Monomer conjugate, Vitamin M-Insulin hexamer conjugate, Vitamin H-Insulin hexamer conjugate, Inulin-Insulin hexamer conjugate, Vitamin M-Insulin Lispro conjugate, Vitamin B1 -Insulin Lispro conjugate, and Inulin- Insulin Lispro conjugate are selected as capable oral-insulin conjugates to interact with Insulin Receptor. In the case of Insulin Receptor, the leucine-rich repeat domain (L1, residues 1-157) and C-terminus of the α-chain (αCT, residues 704-715) are Insulin binding surface[147-150] and they functions as a signaling element to activate its tyrosine kinase and predicted to influence Insulin receptor–Oral insulin conjugate interaction. In the proposed work, none of oral Insulin conjugates shows interaction with αCT while Insulin Monomer - DDM Conjugates (Figure 5a) does not show interaction with L1 domain (Table 8 & Supplementary Figure 8). Insulin Hexamer - DDM Conjugates interacts with ARG86 and ASN34 residues of IR (Figure 5b), while Insulin Lispro - DDM Conjugates interacts with ARG86, ASN90 and ARG114 residues of IR (Figure 5c) which reflects the efficient binding affinity of Insulin Lispro – DDM Conjugates with IR.

In our analysis, fourteen drug delivering agents were screened and its characteristic features for oral delivery of Insulin were examined. Based on the toxicity and conjugation ability with various forms of Insulin, the drug delivering molecules were chosen for developmental studies. From the overall results we nominate Vitamin B1 and Inulin are suitable drug delivering agents because: 1) Vitamin B1 completely satisfy the defined criteria based on bio-text mining; it shows better pharmacophoric and efficient ADME features; 58,59 it is carcinogen but low toxic;[60] as per drug informatics analysis; it shows efficient conjugation with Insulin monomer, Insulin hexamer and Insulin Lispro. Vitamin B1- Insulin hexamer conjugate, and Vitamin B1- Insulin Lispro conjugates shows efficient interaction with IR. Vitamin B1 is a highly recommended molecule for in-vitro and in-vivo studies, 2) Inulin absolutely satisfy the defined criteria based on bio-text mining; it shows superior pharmacophoric and moderate ADME features; 7467,81 it is non-carcinogen, non-mutagen, low toxic, as per drug informatics analysis. It shows efficient conjugation with Insulin monomer, Insulin hexamer, and Insulin Lispro. Inulin-insulin hexamer and Inulin - Insulin Lispro conjugates shows efficient interaction with IR. Inulin is a highly recommended molecule for in-vitro and in-vivo studies.

Moreover, we choose Poly (propylene glycol) is a possible drug delivering agent because it is partially fulfills the defined criteria based on bio-text mining; it shows moderate pharmacophoric and moderate ADME features;[115116122] it is non-carcinogen, non-mutagen, moderately toxic;[126] as per drug informatics analysis. It shows reasonable conjugation with insulin monomer, insulin hexamer, and Insulin Lispro. In earlier studies, propylene glycol showed toxic symptoms after the frequent doses & repeated application when used as a vehicle in medicinal preparations.[124125] Meantime, in drug delivering mechanisms, it eliminates the toxic degradation,[126] and less toxic than the parent substance (Polyethylene glycol),[127] because PEG is a frequently used drug delivering molecule for oral insulin delivery[38]. Consequently, Poly (propylene glycol) may be a possible carrier for further studies.

In another stand, while compare the binding efficiency of various insulin forms; Insulin Lispro (LysB28) shows the competent conjugation with drug delivering molecules and the resultant conjugates shows therapeutically capable interaction with IR than Insulin Monomeric and hexameric form of conjugates.

4. CONCLUSION
The oral bioavailability of insulin is 1%. It may be enhanced by novel carriers that deliver insulin to the site of absorption. In the proposed work, based on the defined criteria 14 drug delivering molecules were filtered from 7 reputed compound databases and it’s Pharmacophoric and ADMET properties were analyzed by Biomedical text mining and drug-informatic tools. The results from the conducted studies concluded that Insulin and Vitamin B1 are considered as novel, safe and proficient oral carriers for Insulin. (Polyethylene glycol) is an optional vehicle for oral delivery of Insulin. Insulin Lispro is the tremendous option for oral delivery than other Insulin forms. Clinical studies are recommended to develop our results.

REFERENCES


Flowchart 1: Screening process of Drug delivering molecules
Figure 1
Pharmacophoric features of Drug delivering molecules are illustrated by Discovery Studio software. Acceptors (Green in color), Donors (Magenta in color) and Hydrophobic Regions (Blue in color) of Pharmacophoric features of all Drug delivering molecules are demonstrated and differentiated by color. C1 - Vitamin B12 (cobalamin); C2 - Vitamin H (Biotin); C3 - Folic acid (Vitamin M / Vitamin B9); C4 - Vitamin B1 (Thiamin); C5 - L-Carnitine (Vitamin BT); C6 - Poly-N-vinylpyrrolidone; C7 - Inulin; C8 - Poly Cysteine; C9 - Chitosan; C10 - Pectin; C11 - Poly(propylene glycol); C12 - Poly(propylene imine); C13 - Poly (lactic-co-glycolic acid); C14 - Deoxycholic acid
Figure 2a
Bioconjugate from Human Insulin Monomer (PDB ID: 1HLS) and Vitamin H is illustrated by Discovery Studio software. Vitamin H conjugated at CYS A11, GLN B4 and HIS B10 aminoacids of Insulin Monomer. The inner figure illustrates the CYS 11 amino acid of Insulin-A chain conjugates with Vitamin H.

Figure 2b
Bioconjugate from Human Insulin Monomer (PDB ID: 1HLS) and Vitamin B1 is illustrated by Discovery Studio software. Vitamin B1 conjugated at CYS A6, CYS A11 and LEU B6 aminoacids of Insulin Monomer. The inner figure illustrates the CYS 6, CYS 11 amino acids of Insulin-A chain and LEU 6 amino acid of Insulin-B chain conjugates with Vitamin B1.
Figure 2c
Bioconjugate from Human Insulin Monomer (PDB ID: 1HLS) and Inulin is illustrated by Discovery Studio software. Inulin conjugated at PHE B1 aminoacid of Insulin Monomer. The inner figure illustrates the PHE 1 amino acid of Insulin-B chain conjugates with Inulin.

Figure 2d
Bioconjugate from Human Insulin Monomer (PDB ID: 1HLS) and Poly (lactic-co-glycolic acid) is illustrated by Discovery Studio software. Poly (lactic-co-glycolic acid) conjugated at PHE B1 and CYS A11 aminoacid of Insulin Monomer. The inner figure illustrates the PHE 1 amino acid of Insulin-B chain conjugates and with Poly (lactic-co-glycolic acid).
Figure 3a
Bioconjugate from Human insulin hexamer (PDB ID: 1AIO) and Vitamin M is illustrated by Discovery Studio software. Vitamin M conjugated at GLY A1 and THR B27 aminoacids of Insulin hexamer. The inner figure illustrates the GLY 1 amino acid of Insulin-A chain conjugates with Vitamin M.

Figure 3b
Bioconjugate from Human insulin hexamer (PDB ID: 1AIO) and Vitamin H is illustrated by Discovery Studio software. Vitamin H conjugated at GLN B4 and ARG B22 aminoacids of Insulin hexamer. The inner figure illustrates the ARG 22 amino acid of Insulin-B chain conjugates with Vitamin H.
Figure 3c
Bioconjugate from Human insulin hexamer (PDB ID: 1AIO) and Inulin is illustrated by Discovery Studio software. Inulin conjugated at LYS B29 aminoacid of Insulin hexamer. The inner figure illustrates the LYS 29 amino acid of Insulin-B chain conjugates with Inulin.

Figure 3d
Bioconjugate from Human insulin hexamer (PDB ID: 1AIO) and Vitamin B1 is illustrated by Discovery Studio software. Vitamin B1 conjugated at LYS B29 aminoacid of Insulin hexamer. The inner figure illustrates the LYS 29 amino acid of Insulin-B chain conjugates with Vitamin B1.
Figure 4a
Bioconjugate from Insulin Lispro (PDB ID: 1LPH) and Vitamin M is illustrated by Discovery Studio software. Vitamin M conjugated at GLY A1, ILE A2, VAL A3, GLU A4, TYR A19, GLN B4, HIS B5, THR B27, LYS B28 and THR B30 aminoacids of Insulin Lispro. The inner figure illustrates the LYS 28 amino acid of Insulin-B chain conjugates with Vitamin M.

Figure 4b
Bioconjugate from Insulin Lispro (PDB ID: 1LPH) and Vitamin B1 is illustrated by Discovery Studio software. Vitamin B1 conjugated at GLU A4 and LYS B28 aminoacids of Insulin Lispro. The inner figure illustrates the LYS 28 amino acid of Insulin-B chain conjugates with Vitamin B1.
Figure 4c
Bioconjugate from Insulin Lispro (PDB ID: 1LPH) and Inulin is illustrated by Discovery Studio software. Inulin conjugated at GLY A1, VAL A3, GLU A4 and LYS B28 aminoacids of Insulin Lispro. The inner figure illustrates the LYS 28 amino acid of Insulin-B chain conjugates with Inulin.

Figure 5a
Interaction results of Oral insulin conjugates (Insulin Monomer (1HLS)- DDM Conjugates) with Insulin Receptor (IR). It does not show any interaction in leucine-rich repeat domain (L1, residues 1-157) and in C-terminus of the α-chain (αCT, residues 704-715). The inner figure shows the interaction in LYS310 aminoacid of IR which is not responsible for initiation of therapeutical effect.
Figure 5b
Interaction results of Oral insulin conjugates (Insulin Hexamer (1AI0)- DDM Conjugates)- DDM Conjugates) with Insulin Receptor (IR). It shows the interaction in ARG86, ASN34 of leucine-rich repeat domain (L1, residues 1-157) and no interaction in C-terminus of the α-chain (αCT, residues 704-715). The inner figure shows the interaction in ASN34 aminoacid of IR which is responsible for initiation of therapeutic effect.

Figure 5c
Interaction results of Oral insulin conjugates (Insulin Lispro (1LPH) - DDM Conjugates) with Insulin Receptor (IR). It shows the interaction in ARG86, ASN90 and ARG114 of leucine-rich repeat domain (L1, residues 1-157) and no interaction in C-terminus of the α-chain (αCT, residues 704-715). The inner figure shows the interaction in ARG86, ARG114 aminoacids of IR which is responsible for initiation of therapeutic effect.
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
<td>Vitamin B12 (cobalamin)</td>
<td>CHEBI ID:30411</td>
<td>Mono-disperse&lt;sup&gt;34&lt;/sup&gt; 1,357&lt;sup&gt;35&lt;/sup&gt; Biocompatible &amp; Biodegradable&lt;sup&gt;36,38&lt;/sup&gt; Oral bioavailability is high&lt;sup&gt;36-39&lt;/sup&gt; Stable&lt;sup&gt;37,38,40,41&lt;/sup&gt; Inert&lt;sup&gt;42&lt;/sup&gt;</td>
<td>No literature</td>
<td>Moderate&lt;sup&gt;50&lt;/sup&gt; Non-toxic&lt;sup&gt;50&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C2</td>
<td>Vitamin H (Biotin)</td>
<td>CHEBI ID:15956</td>
<td>Mono-disperse&lt;sup&gt;45&lt;/sup&gt; 244 Biocompatible &amp; Biodegradable&lt;sup&gt;46,48&lt;/sup&gt; Efficient intestinal transport&lt;sup&gt;46-48&lt;/sup&gt; Stable&lt;sup&gt;46-48&lt;/sup&gt; Inert&lt;sup&gt;49&lt;/sup&gt; Non -toxic&lt;sup&gt;50&lt;/sup&gt;</td>
<td>No literature</td>
<td>Non literature Non -toxic&lt;sup&gt;54-56&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C3</td>
<td>Folic acid (Vitamin M / Vitamin B9)</td>
<td>CHEBI ID:27470</td>
<td>Mono-disperse&lt;sup&gt;31&lt;/sup&gt; 441.40 Biocompatible &amp; Biodegradable&lt;sup&gt;52,54&lt;/sup&gt; Oral bioavailability is high&lt;sup&gt;54,55&lt;/sup&gt; Stable&lt;sup&gt;54,55&lt;/sup&gt; No literature</td>
<td>No literature</td>
<td>Non literature Non -toxic&lt;sup&gt;54-56&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C4</td>
<td>Vitamin B1 (Thiamin)</td>
<td>CHEMBL ID:1547</td>
<td>No literature 265.112 Biocompatible &amp; Biodegradable&lt;sup&gt;57&lt;/sup&gt; Oral bioavailability is high&lt;sup&gt;58&lt;/sup&gt; Stable&lt;sup&gt;58&lt;/sup&gt; Inert&lt;sup&gt;59&lt;/sup&gt; Non-toxic&lt;sup&gt;60&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C5</td>
<td>L-Carnitine (Vitamin BT)</td>
<td>CHEMBL ID:1149</td>
<td>No literature 162.113 Biocompatible &amp; Biodegradable&lt;sup&gt;61,64&lt;/sup&gt; Efficient intestinal transport&lt;sup&gt;61-65&lt;/sup&gt; Stable&lt;sup&gt;61-65&lt;/sup&gt; No literature</td>
<td>No literature</td>
<td>Non literature Non -toxic&lt;sup&gt;68&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C6</td>
<td>Poly-N-vinylpyrrolidone</td>
<td>Pubchem Compound ID: 6917</td>
<td>Mono-disperse&lt;sup&gt;67-68&lt;/sup&gt; 11.141 Biocompatible &amp; Biodegradable&lt;sup&gt;69&lt;/sup&gt; Oral bioavailability is low&lt;sup&gt;70,71&lt;/sup&gt; Stable&lt;sup&gt;72,73&lt;/sup&gt; Inert&lt;sup&gt;74&lt;/sup&gt; Moderately toxic&lt;sup&gt;75&lt;/sup&gt;</td>
<td>No literature</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C7</td>
<td>Inulin</td>
<td>Zinc ID: 12358861</td>
<td>Mono-disperse&lt;sup&gt;76&lt;/sup&gt; 342.297 Biocompatible &amp; Biodegradable&lt;sup&gt;77&lt;/sup&gt; Oral bioavailability is high&lt;sup&gt;78&lt;/sup&gt; Stable&lt;sup&gt;79,81&lt;/sup&gt; Inert&lt;sup&gt;82,83&lt;/sup&gt; Non toxic&lt;sup&gt;84-87&lt;/sup&gt;</td>
<td>No literature</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C8</td>
<td>Poly Cysteine</td>
<td>DrugBank ID: DB000151</td>
<td>Mono-disperse&lt;sup&gt;88&lt;/sup&gt; 121.158&lt;sup&gt;89&lt;/sup&gt; Biocompatible &amp; Biodegradable&lt;sup&gt;90,91&lt;/sup&gt; Hydrophilic; poor intestinal absorption&lt;sup&gt;92,93&lt;/sup&gt; Stable in the presence of the peptidase α-chymotrypsin, increase insulin stability against</td>
<td>No literature</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No.</td>
<td>Name</td>
<td>Identifier</td>
<td>Molecular Weight</td>
<td>Biocompatibility &amp; Biodegradability</td>
<td>Intestinal Permeability</td>
<td>Stability</td>
<td>Toxicity</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----</td>
<td>-----------------------</td>
<td>-----------------</td>
<td>------------------</td>
<td>-------------------------------------</td>
<td>-------------------------</td>
<td>-----------</td>
<td>----------------------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C9</td>
<td>Chitosan</td>
<td>ChemSpider ID: 2342878</td>
<td>Mono-disperse96-98</td>
<td>501.482</td>
<td>Stable96-102</td>
<td>Non-toxic</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C10</td>
<td>Pectin</td>
<td>KEGG ID: C00714</td>
<td>Mono-disperse105</td>
<td>194.139</td>
<td>Stable105</td>
<td>No literature</td>
<td>Non-toxic109</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C11</td>
<td>Poly(propylene glycol)</td>
<td>CHEBI ID:53262</td>
<td>Mono-disperse112,13</td>
<td>76.094</td>
<td>Stable112-116</td>
<td>Inert118-122</td>
<td>Moderately Toxic123-125</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C12</td>
<td>Poly(propylene imine)</td>
<td>CHEBI ID: 53266</td>
<td>Mono-disperse128,129</td>
<td>Molecular weight varies based on length of chain</td>
<td>Biocompatible &amp; Biodegradable130</td>
<td>Lipophilic131</td>
<td>Stable130</td>
<td>Moderately Toxic132</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C13</td>
<td>Poly (lactic-co-glycolic acid)</td>
<td>CHEBI:53493</td>
<td>Mono-disperse133,134</td>
<td>Molecular weight varies based on length of chain</td>
<td>Biocompatible &amp; Biodegradable135-137</td>
<td>Lipophilic138,139</td>
<td>Stable136-139</td>
<td>Inert137</td>
<td>Non-toxic140</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C14</td>
<td>Deoxycholic acid</td>
<td>DrugBank ID: DB03619</td>
<td>Mono-disperse141</td>
<td>392.57</td>
<td>Biocompatible &amp; Biodegradable142,143</td>
<td>Lipophilic144,145</td>
<td>Stable146,145</td>
<td>Not inert144,145</td>
<td>Moderately Toxic146</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Carrier No</td>
<td>Carrier</td>
<td>Number of Acceptor in molecule</td>
<td>Number of Donors in molecule</td>
<td>Number of hydrophobic Region</td>
<td>Number of Pharmacophore principle</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------</td>
<td>----------------------------</td>
<td>--------------------------------</td>
<td>------------------------------</td>
<td>------------------------------</td>
<td>----------------------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C1</td>
<td>Vitamin B12 (cobalamin)</td>
<td>8</td>
<td>7</td>
<td>29</td>
<td>Nil</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C2</td>
<td>Vitamin H (Biotin)</td>
<td>2</td>
<td>2</td>
<td>5</td>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C3</td>
<td><em>Folic acid</em> (Vitamin M / Vitamin B9)</td>
<td>4</td>
<td>4</td>
<td>7</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C4</td>
<td>Vitamin B1 (Thiamin)</td>
<td>1</td>
<td>1</td>
<td>7</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C5</td>
<td>L-Carnitine (Vitamin BT)</td>
<td>1</td>
<td>0</td>
<td>6</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C6</td>
<td>Poly-N-vinylpyrrolidone</td>
<td>1</td>
<td>0</td>
<td>4</td>
<td>Nil</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C7</td>
<td>Inulin</td>
<td>11</td>
<td>0</td>
<td>12</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C8</td>
<td>Poly Cysteine</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C9</td>
<td>Chitosan</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C10</td>
<td>Pectin</td>
<td>3</td>
<td>0</td>
<td>6</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C11</td>
<td>Poly(propylene glycol)</td>
<td>1</td>
<td>0</td>
<td>4</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C12</td>
<td>Poly(propylene imine)</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>Nil</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C13</td>
<td>Poly (lactic-co-glycolic acid)</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C14</td>
<td>Deoxycholic acid</td>
<td>2</td>
<td>0</td>
<td>14</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 3
Physico-chemical properties of drug delivering molecules

<table>
<thead>
<tr>
<th>S. No</th>
<th>Carriers</th>
<th>ADME Descriptors</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Aqueous Solubility</td>
</tr>
<tr>
<td>1</td>
<td>C1</td>
<td>-9.057</td>
</tr>
<tr>
<td>2</td>
<td>C2</td>
<td>-1.432</td>
</tr>
<tr>
<td>3</td>
<td>C3</td>
<td>-3.378</td>
</tr>
<tr>
<td>4</td>
<td>C4</td>
<td>-1.335</td>
</tr>
<tr>
<td>5</td>
<td>C5</td>
<td>2.734</td>
</tr>
<tr>
<td>6</td>
<td>C6</td>
<td>-0.550</td>
</tr>
<tr>
<td>7</td>
<td>C7</td>
<td>0.894</td>
</tr>
<tr>
<td>8</td>
<td>C8</td>
<td>0.335</td>
</tr>
<tr>
<td>9</td>
<td>C9</td>
<td>1.736</td>
</tr>
<tr>
<td>10</td>
<td>C10</td>
<td>1.595</td>
</tr>
<tr>
<td>11</td>
<td>C11</td>
<td>0.907</td>
</tr>
<tr>
<td>12</td>
<td>C12</td>
<td>-0.097</td>
</tr>
<tr>
<td>13</td>
<td>C13</td>
<td>0.768</td>
</tr>
<tr>
<td>14</td>
<td>C14</td>
<td>-4.409</td>
</tr>
</tbody>
</table>

C1 - Vitamin B12 (cobalamin); C2 - Vitamin H (Biotin); C3 - Folic acid (Vitamin M / Vitamin B9); C4 - Vitamin B1 (Thiamin); C5 - L-Carnitine (Vitamin BT); C6 - Poly-N-vinylpyrrolidone; C7; Inulin; C8 - Poly Cysteine; C9 - Chitosan; C10 - Pectin; C11 - Poly(propylene glycol); C12 - Poly(propylene imine); C13 - Poly (lactic-co-glycolic acid); C14 - Deoxycholic acid
<table>
<thead>
<tr>
<th>S. No</th>
<th>Carriers</th>
<th>FDA Rodent Carcinogenicity</th>
<th>Mutagenecity</th>
<th>Rat oral LD50 (g/Kg Body weight)</th>
<th>Rat Maximum tolerated dose (g/Kg Body weight)</th>
<th>Skin Irritancy</th>
<th>Skin sensitization</th>
<th>Aerobic Biodegradability</th>
<th>Toxicity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>C1</td>
<td>Non-carcinogen</td>
<td>Non-mutagen</td>
<td>0.093</td>
<td>0.000</td>
<td>Mild-Irritant</td>
<td>Non-sensitizer</td>
<td>Degradable</td>
<td>High</td>
</tr>
<tr>
<td>2</td>
<td>C2</td>
<td>Non-carcinogen</td>
<td>Non-mutagen</td>
<td>1.109</td>
<td>0.193</td>
<td>Non-Irritant</td>
<td>Non-sensitizer</td>
<td>Degradable</td>
<td>Low</td>
</tr>
<tr>
<td>3</td>
<td>C3</td>
<td>Non-carcinogen</td>
<td>Non-mutagen</td>
<td>2.819</td>
<td>1.391</td>
<td>Non-Irritant</td>
<td>Weak-sensitizer</td>
<td>Non-Degradable</td>
<td>High</td>
</tr>
<tr>
<td>4</td>
<td>C4</td>
<td>Carcinogen</td>
<td>Non-mutagen</td>
<td>1.308</td>
<td>0.097</td>
<td>Non-Irritant</td>
<td>Strong-sensitizer</td>
<td>Non-Degradable</td>
<td>Low</td>
</tr>
<tr>
<td>5</td>
<td>C5</td>
<td>Non-carcinogen</td>
<td>Non-mutagen</td>
<td>1.101</td>
<td>0.175</td>
<td>Mild-Irritant</td>
<td>Non-sensitizer</td>
<td>Degradable</td>
<td>Low</td>
</tr>
<tr>
<td>6</td>
<td>C6</td>
<td>Carcinogen</td>
<td>Non-mutagen</td>
<td>1.634</td>
<td>0.181</td>
<td>Mild-Irritant</td>
<td>Non-sensitizer</td>
<td>Degradable</td>
<td>High</td>
</tr>
<tr>
<td>7</td>
<td>C7</td>
<td>Non-carcinogen</td>
<td>Non-mutagen</td>
<td>20.789</td>
<td>0.000</td>
<td>Mild-Irritant</td>
<td>Non-sensitizer</td>
<td>Degradable</td>
<td>Low</td>
</tr>
<tr>
<td>8</td>
<td>C8</td>
<td>Non-carcinogen</td>
<td>Non-mutagen</td>
<td>0.514</td>
<td>0.653</td>
<td>Non-Irritant</td>
<td>Non-sensitizer</td>
<td>Degradable</td>
<td>Low</td>
</tr>
<tr>
<td>9</td>
<td>C9</td>
<td>Non-carcinogen</td>
<td>Non-mutagen</td>
<td>3.241</td>
<td>0.268</td>
<td>Mild-Irritant</td>
<td>Non-sensitizer</td>
<td>Degradable</td>
<td>High</td>
</tr>
<tr>
<td>10</td>
<td>C10</td>
<td>Non-carcinogen</td>
<td>Non-mutagen</td>
<td>3.576</td>
<td>0.525</td>
<td>Mild-Irritant</td>
<td>Non-sensitizer</td>
<td>Degradable</td>
<td>Low</td>
</tr>
<tr>
<td>11</td>
<td>C11</td>
<td>Non-carcinogen</td>
<td>Non-mutagen</td>
<td>12.098</td>
<td>0.187</td>
<td>Mild-Irritant</td>
<td>Non-sensitizer</td>
<td>Degradable</td>
<td>High</td>
</tr>
<tr>
<td>12</td>
<td>C12</td>
<td>Non-carcinogen</td>
<td>Non-mutagen</td>
<td>0.055</td>
<td>0.089</td>
<td>Mild-Irritant</td>
<td>Non-sensitizer</td>
<td>Degradable</td>
<td>High</td>
</tr>
<tr>
<td>13</td>
<td>C13</td>
<td>Non-carcinogen</td>
<td>Non-mutagen</td>
<td>2.982</td>
<td>0.427</td>
<td>Non-Irritant</td>
<td>Non-sensitizer</td>
<td>Degradable</td>
<td>Low</td>
</tr>
<tr>
<td>14</td>
<td>C14</td>
<td>Non-carcinogen</td>
<td>Non-mutagen</td>
<td>6.358</td>
<td>0.190207</td>
<td>Severe-Irritant Weak-sensitizer</td>
<td>Degradable</td>
<td>High</td>
<td></td>
</tr>
</tbody>
</table>

**FDA** – Food & Drug Administration;  
C1 - Vitamin B12 (cobalamin); C2 - Vitamin H (Biotin); C3 - Folic acid (Vitamin M / Vitamin B9); C4 - Vitamin B1 (Thiamin); C5 - L-Carnitine (Vitamin BT); C6 - Poly-N-vinylpyrrolidone; C7 - Inulin; C8 - Poly Cysteine; C9 – Chitosan; C10 – Pectin; C11 - Poly(propylene glycol); C12 - Poly(propylene imine); C13 - Poly (lactic-co-glycolic acid); C14 - Deoxycholic acid
<table>
<thead>
<tr>
<th>S.No</th>
<th>Conjugate</th>
<th>Binding site (LYS B29)</th>
<th>LIP DOCK SCORE</th>
<th>Binding site (PHE B1)</th>
<th>LIP DOCK SCORE</th>
<th>Other Binding site</th>
<th>LIP DOCK SCORE</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
<td>Vitamin B12</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>C2</td>
<td>Vitamin H</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>CYS A11, HIS B10, GLN B4</td>
<td>86.2835</td>
</tr>
<tr>
<td>C3</td>
<td>Vitamin M</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>C4</td>
<td>Vitamin B1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>CYS A6, CYS A11, LEU B6</td>
<td>94.1144</td>
</tr>
<tr>
<td>C5</td>
<td>Vitamin BT</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>LEU A13, VAL B2</td>
<td>68.0592</td>
</tr>
<tr>
<td>C6</td>
<td>poly-N-vinylpyrrolidone</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>C7</td>
<td>Inulin</td>
<td>-</td>
<td>-</td>
<td>PHE B1</td>
<td>117.663</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>C8</td>
<td>Poly Cysteine</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>GLN B4</td>
<td>53.1597</td>
</tr>
<tr>
<td>C9</td>
<td>Chitosan</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>CYS A11, VAL B2, ASN B3</td>
<td>74.4251</td>
</tr>
<tr>
<td>C10</td>
<td>PECTIN</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>CYS A11</td>
<td>63.0902</td>
</tr>
<tr>
<td>C11</td>
<td>POLY (PROPYLENE GLYCOL)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td>76.0238</td>
</tr>
<tr>
<td>C12</td>
<td>poly(propylene imine)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>VAL B2</td>
<td>76.0238</td>
</tr>
<tr>
<td>C13</td>
<td>Poly (lactic-co-glycolic acid)</td>
<td>-</td>
<td>-</td>
<td>PHE B1</td>
<td>68.6737</td>
<td>CYS A11</td>
<td>68.6737</td>
</tr>
<tr>
<td>C14</td>
<td>Deoxycholic acid</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>S.No</td>
<td>Conjugate</td>
<td>Binding site (LYS B29)</td>
<td>LIP DOCK SCORE -</td>
<td>Binding site (PHE B1)</td>
<td>LIP DOCK SCORE -</td>
<td>Other Binding site</td>
<td>LIP DOCK SCORE</td>
</tr>
<tr>
<td>------</td>
<td>---------------------------</td>
<td>------------------------</td>
<td>-----------------</td>
<td>-----------------------</td>
<td>-----------------</td>
<td>--------------------</td>
<td>----------------</td>
</tr>
<tr>
<td>C1</td>
<td>Vit B12</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>C2</td>
<td>Vit H</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>GLN B4 ARG B22</td>
<td>103.231</td>
</tr>
<tr>
<td>C3</td>
<td>Vit M</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>GLY A1 THR B27</td>
<td>114.324</td>
</tr>
<tr>
<td>C4</td>
<td>Vit B1</td>
<td>LYS B29</td>
<td>79.8834</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>C5</td>
<td>Vit BT</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>ARG B22</td>
<td>84.7767</td>
</tr>
<tr>
<td>C6</td>
<td>poly-N-vinylpyrrolidone</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>C7</td>
<td>Inulin</td>
<td>LYS B29</td>
<td>94.3543</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>C8</td>
<td>Poly Cysteine</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>TYR B16 GLU B21</td>
<td>58.5629</td>
</tr>
<tr>
<td>C9</td>
<td>Chitosan</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>TYR B16 TYR B26</td>
<td>90.016</td>
</tr>
<tr>
<td>C10</td>
<td>PECTIN</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>GLN B4</td>
<td>91.0549</td>
</tr>
<tr>
<td>C11</td>
<td>POLY (PROPYLENE GLYCOL)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>GLU B21 ARG B22</td>
<td>71.5126</td>
</tr>
<tr>
<td>C12</td>
<td>poly(propylene imine)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>C13</td>
<td>Poly (lactic-co-glycolic acid)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>GLY B20 ARG B22</td>
<td>82.2602</td>
</tr>
<tr>
<td>C14</td>
<td>Deoxycholic acid</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>GLY A1 ILE A2</td>
<td>60.763</td>
</tr>
</tbody>
</table>
Table 7
Conjugation results of Insulin Lispro (PDB ID: 1LPH), with all listed drug delivering molecules individually

<table>
<thead>
<tr>
<th>S.No</th>
<th>Conjugate</th>
<th>Binding site (LYS B28)</th>
<th>LIP DOCK SCORE</th>
<th>Binding site (PHE B1)</th>
<th>LIP DOCK SCORE</th>
<th>Other Binding site</th>
<th>LIP DOCK SCORE</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
<td>Vit B12</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>C2</td>
<td>Vit H</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>C3</td>
<td>Vit M</td>
<td>LYS B28</td>
<td>131.57</td>
<td>-</td>
<td>-</td>
<td>GLY A1, ILE A2, VAL A3, GLU A4, TYR A19, GLN B4 HIS B5 THR B27, THR B30</td>
<td>131.57</td>
</tr>
<tr>
<td>C4</td>
<td>Vit B1</td>
<td>LYS B28</td>
<td>89.8971</td>
<td>-</td>
<td>-</td>
<td>GLU A4</td>
<td>89.8971</td>
</tr>
<tr>
<td>C5</td>
<td>Vit BT</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>GLU A4, GLN B4</td>
<td>58.6757</td>
</tr>
<tr>
<td>C6</td>
<td>poly-N-vinylpyrrolidone</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>C7</td>
<td>Inulin</td>
<td>LYS B28</td>
<td>76.2195</td>
<td>-</td>
<td>-</td>
<td>GLY A1, VAL A3, GLU A4</td>
<td>76.2195</td>
</tr>
<tr>
<td>C8</td>
<td>Poly Cysteine</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>GLN B4, GLU A4</td>
<td>48.858</td>
</tr>
<tr>
<td>C9</td>
<td>Chitosan</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>GLY A1, ILE A2, GLU A4 THR B27</td>
<td>67.1639</td>
</tr>
<tr>
<td>C10</td>
<td>PECTIN</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>GLY A1, GLU A4</td>
<td>73.2077</td>
</tr>
<tr>
<td>C11</td>
<td>POLY (PROPYLENE GLYCOL)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>GLY A1</td>
<td>57.6819</td>
</tr>
<tr>
<td>C12</td>
<td>poly(propylene imine)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>C13</td>
<td>Poly (lactic-co-glycolic acid)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>GLY A1, ILE A2, GLU A4, VAL A3, THR B27, GLN 44</td>
<td>64.8442</td>
</tr>
<tr>
<td>C14</td>
<td>Deoxycholic acid</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 8
Interaction results of Oral insulin conjugates (OIC) with Insulin Receptor (IR)

<table>
<thead>
<tr>
<th>S. No</th>
<th>Conjugates</th>
<th>Interaction with leucine-rich repeat domain (L1, residues 1-157)</th>
<th>Interaction with C-terminus of the α-chain (αCT, residues 704-715)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Insulin Monomer (1HLS)- DDM Conjugates</td>
<td>No Interaction</td>
<td>No Interaction</td>
</tr>
<tr>
<td>2</td>
<td>Insulin Hexamer (1AIQ)- DDM Conjugates</td>
<td>ARG86 ASN34</td>
<td>No Interaction</td>
</tr>
<tr>
<td>3</td>
<td>Insulin Lispro (1LPH) - DDM Conjugates</td>
<td>ARG86 ASN90 ARG114</td>
<td>No Interaction</td>
</tr>
</tbody>
</table>
Supplementary figure 1
Pharmacophoric principles of Drug delivering molecules are illustrated by Discovery Studio software. Acceptors (Green in color), Donors (Magenta in color) of Pharmacophoric principles of all Drug delivering molecules are demonstrated and differentiated by color. C1 - Vitamin B12 (cobalamin); C2 - Vitamin H (Biotin); C3 - Folic acid (Vitamin M / Vitamin B9); C4 - Vitamin B1 (Thiamin); C5 - L-Carnitine (Vitamin BT); C6 - Poly-N-vinylpyrrolidone; C7 - Inulin; C8 - Poly Cysteine; C9 - Chitosan; C10 - Pectin; C11 - Poly(propylene glycol); C12 - Poly(propylene imine); C13 - Poly (lactic-co-glycolic acid); C14 - Deoxycholic acid

Supplementary Figure (1a): C1 - Vitamin B12 (cobalamin)
Number of Pharmacophoric principles – Nil

Supplementary Figure (1b): C2 - Vitamin H (Biotin)
Number of Pharmacophoric principles – 6
Supplementary Figure (1c): C3 - Folic acid (Vitamin M / Vitamin B9)
Number of Pharmacophoric principles -10
Supplementary Figure (1d): C4 - Vitamin B1 (Thiamin)
Number of Pharmacophoric principles = 10
Supplementary Figure (1e): C5 - L-Carnitine (Vitamin BT)
Number of Pharmacophoric principles – 2

Supplementary Figure (1f): C6- Poly-N-vinylpyrrolidone
Number of Pharmacophoric principles – Nil
Supplementary Figure (1g): C7- Inulin
Number of Pharmacophoric principles – 10
Supplementary Figure (1h): C8- Poly Cysteine
Number of Pharmacophoric principles - 7
Supplementary Figure (1i): C9 – Chitosan
Number of Pharmacophoric principles – 10
Supplementary Figure (1j): C10 – Pectin
Number of Pharmacophoric principles – 10
Supplementary Figure (1k): C11 – Poly(propylene glycol)
Number of Pharmacophoric principles –4

Supplementary Figure (1l): C12 – Poly(propylene imine)
Number of Pharmacophoric principles –Nil
Supplementary Figure (1m): C13 – Poly (lactic-co-glycolic acid)
Number of Pharmacophoric principles – 8
Supplementary Figure (1n): C14 – Deoxycholic acid
Number of Pharmacophoric principles – 7
Supplementary figure 2
ADME descriptors of Drug delivering molecules by Discovery Studio; C1 - Vitamin B12 (cobalamin); C2 - Vitamin H (Biotin); C3 - Folic acid (Vitamin M / Vitamin B9); C4 - Vitamin B1 (Thiamin); C5 - L-Carnitine (Vitamin BT); C6 - Poly-N-vinylpyrrolidone; C7 - Inulin; C8 - Poly Cysteine; C9 - Chitosan; C10 - Pectin; C11 - Poly(propylene glycol); C12 - Poly(propylene imine); C13 - Poly (lactic-co-glycolic acid); C14 - Deoxycholic acid

Supplementary figure 2(a)- C1 - Vitamin B12

Supplementary figure 2(b)- C2 - Vitamin H (Biotin)
**Supplementary figure 2(c) - C3 - Folic acid (Vitamin M / Vitamin B9)**

**Supplementary figure 2(d) - C4 - Vitamin B1 (Thiamin)**
Supplementary figure 2(e) - C5 - L-Carnitine (Vitamin BT)

Supplementary figure 2(f) - C6 - Poly-N-vinylpyrrolidone
Supplementary figure 2(g) - C7 - Inulin

Supplementary figure 2(h) - C8 - Poly Cysteine
Supplementary figure 2(i) - C9 – Chitosan

Supplementary figure 2(j) - C10 – Pectin
Supplementary figure 2(k) - C11 - Poly(propylene glycol)

Supplementary figure 2(l) - C12 - Poly(propylene imine)
Supplementary figure 2(m) - C13 - Poly (lactic-co-glycolic acid)

Supplementary figure 2(n) - C14 - Deoxycholic acid
Supplementary figure 3
Toxicity studies for Drug delivering molecules by Discovery Studio; C1 - Vitamin B12 (cobalamin); C2 - Vitamin H (Biotin); C3 - Folic acid (Vitamin M / Vitamin B9); C4 - Vitamin B1 (Thiamin); C5 - L-Carnitine (Vitamin BT); C6 - Poly-N-vinylpyrrolidone; C7 - Inulin; C8 - Poly Cysteine; C9 - Chitosan; C10 - Pectin; C11 - Poly(propylene glycol); C12 - Poly(propylene imine); C13 - Poly (lactic-co-glycolic acid); C14 - Deoxycholic acid

Supplementary figure 3(a)- C1 - Vitamin B12 (cobalamin)

Model Prediction
Prediction: Non-Carcinogen
Probability: 0.230
Enrichment: 0.717
Bayesian Score: -1.075

Prediction: Positive if the Bayesian score is above the estimated best cutoff value.
Probability: The estimated probability that the sample is in the category.
Enrichment: An estimate of enrichment, that is, the increased likelihood (versus random) of this sample being in the category.
Bayesian Score: The standard Laplacian-modified Bayesian score.

Model Prediction
Prediction: Non-Mutagen
Probability: 0.029
Enrichment: 0.053
Bayesian Score: -21.617

Prediction: Positive if the Bayesian score is above the estimated best cutoff value.
Probability: The estimated probability that the sample is in the category.
Enrichment: An estimate of enrichment, that is, the increased likelihood (versus random) of this sample being in the category.
Bayesian Score: The standard Laplacian-modified Bayesian score.
Rat oral LD50 (g/Kg Body weight)

C_{63}H_{88}CoN_{14}O_{14}P
Molecular Weight: 1355.36518
ALogP: -0.442
Rotatable Bonds: 16
Acceptors: 16
Donors: 9

Model Prediction
Prediction: 0.093
Unit: g/kg_body_weight

Rat Maximum tolerated dose (g/Kg Body weight)

C_{63}H_{88}CoN_{14}O_{14}P
Molecular Weight: 1355.36518
ALogP: -0.442
Rotatable Bonds: 16
Acceptors: 15
Donors: 9

Model Prediction
Prediction: 0.000
Unit: g/kg_body_weight
Skin Irritancy

\[ C_{63}H_{92}CoN_{4}O_{14}P \]
Molecular Weight: 1355.36518
ALogP: -0.442
Rotatable Bonds: 16
Acceptors: 16
Donors: 9

Model Prediction
Prediction: Mild
Probability: 0.071
Enrichment: 0.194
Bayesian Score: -10.755

Prediction: Positive if the Bayesian score is above the estimated best cutoff value.
Probability: The estimated probability that the sample is in the category.
Enrichment: An estimate of enrichment, that is, the increased likelihood (versus random) of this sample being in the category.
Bayesian Score: The standard Laplacian-modified Bayesian score.

Model Prediction
Prediction: Non-Sensitizer
Probability: 0.502
Enrichment: 0.731
Bayesian Score: -4.826

Prediction: Positive if the Bayesian score is above the estimated best cutoff value.
Probability: The estimated probability that the sample is in the category.
Enrichment: An estimate of enrichment, that is, the increased likelihood (versus random) of this sample being in the category.
Bayesian Score: The standard Laplacian-modified Bayesian score.
Aerobic Biodegradability

\[
\text{C}_{35}\text{H}_{88}\text{CoN}_{14}\text{O}_{14}\text{P}
\]
Molecular Weight: 1355.36518
ALogP: -0.442
Rotatable Bonds: 16
Acceptors: 16
Donors: 9

**Model Prediction**

*Prediction: Degradable*

Probability: 0.634
Enrichment: 1.454
Bayesian Score: 3.091

*Prediction: Positive if the Bayesian score is above the estimated best cutoff value.*
*Probability: The estimated probability that the sample is in the category.*
*Enrichment: An estimate of enrichment, that is, the increased likelihood (versus random) of this sample being in the category.*
*Bayesian Score: The standard Laplacian-modified Bayesian score.*
Supplementary figure 3(b) - C2 - Vitamin H (Biotin)

**FDA Rodent Carcinogenicity**

\[
\text{C}_{10}H_{19}N_2O_3S \\
\text{Molecular Weight: 244.31064} \\
\text{ALogP: 0.67} \\
\text{Rotatable Bonds: 5} \\
\text{Acceptors: 4} \\
\text{Donors: 3}
\]

**Mutagenecity**

\[
\text{C}_{10}H_{19}N_2O_3S \\
\text{Molecular Weight: 244.31064} \\
\text{ALogP: 0.67} \\
\text{Rotatable Bonds: 5} \\
\text{Acceptors: 4} \\
\text{Donors: 3}
\]

### Model Prediction

**Prediction: Non-Carcinogen**

- **Probability**: 0.435
- **Enrichment**: 0.845
- **Bayesian Score**: -2.888

Predictions:
- Positive if the Bayesian score is above the estimated best cutoff value.
- Probability: The estimated probability that the sample is in the category.
- Enrichment: An estimate of enrichment, that is, the increased likelihood (versus random) of this sample being in the category.
- Bayesian Score: The standard Laplacian-modified Bayesian score.

**Prediction: Non-Mutagen**

- **Probability**: 0.002
- **Enrichment**: 0.004
- **Bayesian Score**: -29.037

Predictions:
- Positive if the Bayesian score is above the estimated best cutoff value.
- Probability: The estimated probability that the sample is in the category.
- Enrichment: An estimate of enrichment, that is, the increased likelihood (versus random) of this sample being in the category.
- Bayesian Score: The standard Laplacian-modified Bayesian score.
Rat oral LD50 (g/Kg Body weight)  

\[ \text{C}_{10}\text{H}_{18}\text{N}_{2}\text{O}_{3}\text{S} \]  
Molecular Weight: 244.31064  
ALogP: 0.67  
Rotatable Bonds: 5  
Acceptors: 4  
Donors: 3

Model Prediction  
Prediction: 1.109  
Unit: g/kg_body_weight

Rat Maximum tolerated dose (g/Kg Body weight)

\[ \text{C}_{10}\text{H}_{18}\text{N}_{2}\text{O}_{3}\text{S} \]  
Molecular Weight: 244.31064  
ALogP: 0.67  
Rotatable Bonds: 5  
Acceptors: 4  
Donors: 3

Model Prediction  
Prediction: 0.193  
Unit: g/kg_body_weight
Skin Irritancy

Model Prediction
Prediction: Non-Irritant
Probability: 0.971
Enrichment: 1.054
Bayesian Score: -0.769

Prediction: Positive if the Bayesian score is above the estimated best cutoff value.
Probability: The estimated probability that the sample is in the category.
Enrichment: An estimate of enrichment, that is, the increased likelihood (versus random) of this sample being in the category.
Bayesian Score: The standard Laplacian-modified Bayesian score.

Model Prediction
Prediction: Non-Sensitizer
Probability: 0.283
Enrichment: 0.412
Bayesian Score: -7.644

Prediction: Positive if the Bayesian score is above the estimated best cutoff value.
Probability: The estimated probability that the sample is in the category.
Enrichment: An estimate of enrichment, that is, the increased likelihood (versus random) of this sample being in the category.
Bayesian Score: The standard Laplacian-modified Bayesian score.
Aerobic Biodegradability

C_{10}H_{18}N_{2}O_{3}S
Molecular Weight: 244.31064
ALogP: 0.67
Rotatable Bonds: 5
Acceptors: 4
Donors: 3

Model Prediction
Prediction: Degradable
Probability: 0.759
Enrichment: 1.739
Bayesian Score: 6.104

Prediction: Positive if the Bayesian score is above the estimated best cutoff value.
Probability: The estimated probability that the sample is in the category.
Enrichment: An estimate of enrichment, that is, the increased likelihood (versus random) of this sample being in the category.
Bayesian Score: The standard Laplacian-modified Bayesian score.
Supplementary figure 3(c)- C3 - Folic acid (Vitamin M / Vitamin B9)

**Model Prediction**

**Prediction:** Non-Carcinogen  
**Probability:** 0.210  
**Enrichment:** 0.655  
**Bayesian Score:** -4.767  

Prediction: Positive if the Bayesian score is above the estimated best cutoff value.  
Probability: The estimated probability that the sample is in the category.  
Enrichment: An estimate of enrichment, that is, the increased likelihood (versus random) of this sample being in the category.  
Bayesian Score: The standard Laplacian-modified Bayesian score.

**Model Prediction**

**Prediction:** Non-Mutagen  
**Probability:** 0.000  
**Enrichment:** 0.000  
**Bayesian Score:** -60.476  

Prediction: Positive if the Bayesian score is above the estimated best cutoff value.  
Probability: The estimated probability that the sample is in the category.  
Enrichment: An estimate of enrichment, that is, the increased likelihood (versus random) of this sample being in the category.  
Bayesian Score: The standard Laplacian-modified Bayesian score.
**Rat oral LD50 (g/Kg Body weight)**

![Chemical Structure 1]

- Chemical Formula: C_{25}H_{26}N_{2}O_{2}
- Molecular Weight: 441.39745
- ALogP: -0.232
- Rotatable Bonds: 9
- Acceptors: 11
- Donors: 6

**Model Prediction**

- Prediction: 2.819
- Unit: g/kg_body_weight

---

**Rat Maximum tolerated dose (g/Kg Body weight)**

![Chemical Structure 2]

- Chemical Formula: C_{19}H_{16}N_{2}O_{8}
- Molecular Weight: 441.39745
- ALogP: -0.232
- Rotatable Bonds: 9
- Acceptors: 11
- Donors: 6

**Model Prediction**

- Prediction: 1.391
- Unit: g/kg_body_weight
Skin Irritancy

\[
\text{C}_{19}\text{H}_{17}\text{N}_{4}\text{O}_{5}
\]
Molecular Weight: 441.39745
ALogP: -0.232
Rotatable Bonds: 9
Acceptors: 11
Donors: 6

Skin sensitization

\[
\text{C}_{19}\text{H}_{17}\text{N}_{4}\text{O}_{5}
\]
Molecular Weight: 441.39745
ALogP: -0.232
Rotatable Bonds: 9
Acceptors: 11
Donors: 6

Model Prediction
Prediction: Non-Irritant
Probability: 0.942
Enrichment: 1.023
Bayesian Score: -1.988

Model Prediction
Prediction: Weak-Sensitizer
Probability: 0.602
Enrichment: 1.035
Bayesian Score: -2.527

Prediction: Positive if the Bayesian score is above the estimated best cutoff value.
Probability: The estimated probability that the sample is in the category.
Enrichment: An estimate of enrichment, that is, the increased likelihood (versus random) of this sample being in the category.
Bayesian Score: The standard Laplacian-modified Bayesian score.
Aerobic Biodegradability

C_{19}H_{19}N_{8}O_{6}
Molecular Weight: 441.39745
ALogP: -0.232
Rotatable Bonds: 9
Acceptors: 11
Donors: 6

**Model Prediction**

**Prediction: Non-Degradable**
Probability: 0.267
Enrichment: 0.612
Bayesian Score: -5.103

**Prediction:** Positive if the Bayesian score is above the estimated best cutoff value.
**Probability:** The estimated probability that the sample is in the category.
**Enrichment:** An estimate of enrichment, that is, the increased likelihood (versus random) of this sample being in the category.
**Bayesian Score:** The standard Laplacian-modified Bayesian score.
Supplementary figure 3(d) - C4 - Vitamin B1 (Thiamin)

**FDA Rodent Carcinogenicity**

C₄H₁₂N₄OS
Molecular Weight: 265.35458
ALogP: 1.048
Rotatable Bonds: 4
Acceptors: 4
Donors: 2

**Mutagenecity**

C₄H₁₂N₄OS
Molecular Weight: 265.35458
ALogP: 1.048
Rotatable Bonds: 4
Acceptors: 4
Donors: 2

<table>
<thead>
<tr>
<th><strong>Model Prediction</strong></th>
<th><strong>Model Prediction</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Prediction:</strong> Carcinogen</td>
<td><strong>Prediction:</strong> Non-Mutagen</td>
</tr>
<tr>
<td>Probability: 0.239</td>
<td>Probability: 0.274</td>
</tr>
<tr>
<td>Enrichment: 0.747</td>
<td>Enrichment: 0.490</td>
</tr>
<tr>
<td>Bayesian Score: -0.319</td>
<td>Bayesian Score: -12.406</td>
</tr>
</tbody>
</table>

Prediction: Positive if the Bayesian score is above the estimated best cutoff value.
Probability: The estimated probability that the sample is in the category.
Enrichment: An estimate of enrichment, that is, the increased likelihood (versus random) of this sample being in the category.
Bayesian Score: The standard Laplacian-modified Bayesian score.
Rat oral LD50 (g/Kg Body weight)

\[
\text{C}_{12}\text{H}_{17}\text{N}_4\text{OS} \\
\text{Molecular Weight: 265.36458} \\
\text{ALogP: 1.048} \\
\text{Rotatable Bonds: 4} \\
\text{Acceptors: 4} \\
\text{Donors: 2}
\]

Model Prediction
Prediction: 1.308
Unit: g/kg_body_weight

Rat Maximum tolerated dose (g/Kg Body weight)

\[
\text{C}_{12}\text{H}_{17}\text{N}_4\text{OS} \\
\text{Molecular Weight: 265.36458} \\
\text{ALogP: 1.048} \\
\text{Rotatable Bonds: 4} \\
\text{Acceptors: 4} \\
\text{Donors: 2}
\]

Model Prediction
Prediction: 0.097
Unit: g/kg_body_weight
Skin Irritancy

**Model Prediction**

**Prediction:** Non-Irritant  
**Probability:** 0.963  
**Enrichment:** 1.046  
**Bayesian Score:** -1.251

Prediction: Positive if the Bayesian score is above the estimated best cutoff value.  
Probability: The estimated probability that the sample is in the category.  
Enrichment: An estimate of enrichment, that is, the increased likelihood (versus random) of this sample being in the category.  
Bayesian Score: The standard Laplacian-modified Bayesian score.

**Model Prediction**

**Prediction:** Strong-Sensitizer  
**Probability:** 0.920  
**Enrichment:** 1.187  
**Bayesian Score:** 1.059

Prediction: Positive if the Bayesian score is above the estimated best cutoff value.  
Probability: The estimated probability that the sample is in the category.  
Enrichment: An estimate of enrichment, that is, the increased likelihood (versus random) of this sample being in the category.  
Bayesian Score: The standard Laplacian-modified Bayesian score.
Aerobic Biodegradability

C_{12}H_{17}N_{4}OS  
Molecular Weight: 265.35458  
 ALOGP: 1.048  
Rotatable Bonds: 4  
Acceptors: 4  
Donors: 2

**Model Prediction**  
Prediction: Non-Degradable  
Probability: 0.208  
Enrichment: 0.478  
Bayesian Score: -6.865

Prediction: Positive if the Bayesian score is above the estimated best cutoff value.  
Probability: The estimated probability that the sample is in the category.  
Enrichment: An estimate of enrichment, that is, the increased likelihood (versus random) of this sample being in the category.  
Bayesian Score: The standard Laplacian-modified Bayesian score.
Supplementary figure 3(e): C5 - L-Carnitine (Vitamin BT)

**Model Prediction**

**Prediction:** Non-Carcinogen

**Probability:** 0.218

**Enrichment:** 0.681

**Bayesian Score:** -2.327

*Prediction:* Positive if the Bayesian score is above the estimated best cutoff value.
*Probability:* The estimated probability that the sample is in the category.
*Enrichment:* An estimate of enrichment, that is, the increased likelihood (versus random) of this sample being in the category.
*Bayesian Score:* The standard Laplacian-modified Bayesian score.

---

**Model Prediction**

**Prediction:** Non-Mutagen

**Probability:** 0.543

**Enrichment:** 0.972

**Bayesian Score:** -6.539

*Prediction:* Positive if the Bayesian score is above the estimated best cutoff value.
*Probability:* The estimated probability that the sample is in the category.
*Enrichment:* An estimate of enrichment, that is, the increased likelihood (versus random) of this sample being in the category.
*Bayesian Score:* The standard Laplacian-modified Bayesian score.

---

**Molecular Structure**

- **C₆H₁₄NO₃**
- **Molecular Weight:** 161.19889
- **ALogP:** -3.29
- **Rotatable Bonds:** 4
- **Acceptors:** 3
- **Donors:** 1

---

**Molecular Structure**

- **C₆H₁₄NO₃**
- **Molecular Weight:** 161.19889
- **ALogP:** -3.29
- **Rotatable Bonds:** 4
- **Acceptors:** 3
- **Donors:** 1
Rat oral LD50 (g/Kg Body weight)

C_7H_15NO_3
Molecular Weight: 161.19889
ALogP: -3.29
Rotatable Bonds: 4
Acceptors: 3
Donors: 1

Model Prediction
Prediction: 1.101
Unit: g/kg_body_weight

Rat Maximum tolerated dose (g/Kg Body weight)

C_7H_15NO_3
Molecular Weight: 161.19889
ALogP: -3.29
Rotatable Bonds: 4
Acceptors: 3
Donors: 1

Model Prediction
Prediction: 0.175
Unit: g/kg_body_weight
Skin Irritancy

\[
\begin{align*}
\text{Chemical Structure:} & \quad \text{C}_7\text{H}_{15}\text{NO}_3 \\
\text{Molecular Weight:} & \quad 161.19889 \\
\text{ALogP:} & \quad -3.29 \\
\text{Rotatable Bonds:} & \quad 4 \\
\text{Acceptors:} & \quad 3 \\
\text{Donors:} & \quad 1
\end{align*}
\]

**Model Prediction**

**Prediction:** Mild

**Probability:** 0.170

**Enrichment:** 0.462

**Bayesian Score:** -6.587

*Prediction:* Positive if the Bayesian score is above the estimated best cutoff value.

*Probability:* The estimated probability that the sample is in the category.

*Enrichment:* An estimate of enrichment, that is, the increased likelihood (versus random) of this sample being in the category.

*Bayesian Score:* The standard Laplacian-modified Bayesian score.

Skin sensitization

\[
\begin{align*}
\text{Chemical Structure:} & \quad \text{C}_7\text{H}_{15}\text{NO}_3 \\
\text{Molecular Weight:} & \quad 161.19889 \\
\text{ALogP:} & \quad -3.29 \\
\text{Rotatable Bonds:} & \quad 4 \\
\text{Acceptors:} & \quad 3 \\
\text{Donors:} & \quad 1
\end{align*}
\]

**Model Prediction**

**Prediction:** Non-Sensitizer

**Probability:** 0.636

**Enrichment:** 0.926

**Bayesian Score:** -2.950

*Prediction:* Positive if the Bayesian score is above the estimated best cutoff value.

*Probability:* The estimated probability that the sample is in the category.

*Enrichment:* An estimate of enrichment, that is, the increased likelihood (versus random) of this sample being in the category.

*Bayesian Score:* The standard Laplacian-modified Bayesian score.
Aerobic Biodegradability

Chemical Structure

C_{15}H_{19}NO_3
Molecular Weight: 161.19889
 ALOGP: -3.29
 Rotatable Bonds: 4
 Acceptors: 3
 Donors: 1

Model Prediction
Prediction: Degradable
Probability: 0.621
Enrichment: 1.422
Bayesian Score: 2.788

Prediction: Positive if the Bayesian score is above the estimated best cutoff value.
Probability: The estimated probability that the sample is in the category.
Enrichment: An estimate of enrichment, that is, the increased likelihood (versus random) of this sample being in the category.
Bayesian Score: The standard Laplacian-modified Bayesian score.
Supplementary figure 3(f) - C6 - Poly-N-vinylpyrrolidone

**FDA Rodent Carcinogenicity**

![Chemical structure](image)

C₆H₉NO₂
Molecular Weight: 111.14176
ALogP: 6.1e-002
Rotatable Bonds: 2
Acceptors: 1
Donors: 0

**Mutagenecity**

![Chemical structure](image)

C₆H₉NO₂
Molecular Weight: 111.14176
ALogP: 6.1e-002
Rotatable Bonds: 2
Acceptors: 1
Donors: 0

**Model Prediction**

**Prediction:** Carcinogen
**Probability:** 0.256
**Enrichment:** 0.798
**Bayesian Score:** 0.593

**Prediction:** Positive if the Bayesian score is above the estimated best cutoff value.
**Probability:** The estimated probability that the sample is in the category.
**Enrichment:** An estimate of enrichment, that is, the increased likelihood (versus random) of this sample being in the category.
**Bayesian Score:** The standard Laplacian-modified Bayesian score.

**Model Prediction**

**Prediction:** Non-Mutagen
**Probability:** 0.494
**Enrichment:** 0.885
**Bayesian Score:** -7.677

**Prediction:** Positive if the Bayesian score is above the estimated best cutoff value.
**Probability:** The estimated probability that the sample is in the category.
**Enrichment:** An estimate of enrichment, that is, the increased likelihood (versus random) of this sample being in the category.
**Bayesian Score:** The standard Laplacian-modified Bayesian score.
Rat oral LD50 (g/Kg Body weight)  
Rat Maximum tolerated dose (g/Kg Body weight)

\[ \text{C}_6\text{H}_9\text{NO} \]
Molecular Weight: 111.14176  
AlogP: 6.1e-002  
Rotatable Bonds: 2  
Acceptors: 1  
Donors: 0

**Model Prediction**  
Prediction: 1.634  
Unit: g/kg_body_weight

\[ \text{C}_6\text{H}_9\text{NO} \]
Molecular Weight: 111.14176  
AlogP: 6.1e-002  
Rotatable Bonds: 2  
Acceptors: 1  
Donors: 0

**Model Prediction**  
Prediction: 0.181  
Unit: g/kg_body_weight
Skin Irritancy

C₆H₉NO
Molecular Weight: 111.14176
ALogP: 6.1e-002
Rotatable Bonds: 2
Acceptors: 1
Donors: 0

Model Prediction
Prediction: Mild
Probability: 0.311
Enrichment: 0.844
Bayesian Score: -2.583

Prediction: Positive if the Bayesian score is above the estimated best cutoff value.
Probability: The estimated probability that the sample is in the category.
Enrichment: An estimate of enrichment, that is, the increased likelihood (versus random) of this sample being in the category.
Bayesian Score: The standard Laplacian-modified Bayesian score.

Skin sensitization

C₆H₉⁺NO
Molecular Weight: 111.14176
ALogP: 6.1e-002
Rotatable Bonds: 2
Acceptors: 1
Donors: 0

Model Prediction
Prediction: Non-Sensitizer
Probability: 0.337
Enrichment: 0.492
Bayesian Score: -6.913

Prediction: Positive if the Bayesian score is above the estimated best cutoff value.
Probability: The estimated probability that the sample is in the category.
Enrichment: An estimate of enrichment, that is, the increased likelihood (versus random) of this sample being in the category.
Bayesian Score: The standard Laplacian-modified Bayesian score.
Aerobic Biodegradability

\[ \text{C}_9\text{H}_{15}\text{N}_2\text{O}_2 \]

Molecular Weight: 111.14176
ALogP: 6.1e-002
Rotatable Bonds: 2
Acceptors: 1
Donors: 0

**Model Prediction**

*Prediction: Degradable*

Probability: 0.703
Enrichment: 1.610
Bayesian Score: 4.656

*Prediction: Positive if the Bayesian score is above the estimated best cutoff value.*
*Probability: The estimated probability that the sample is in the category.*
*Enrichment: An estimate of enrichment, that is, the increased likelihood (versus random) of this sample being in the category.*
*Bayesian Score: The standard Laplacian-modified Bayesian score.*
Supplementary figure 3(g): C7; Inulin

**FDA Rodent Carcinogenicity**

![C7 Inulin](image)

- **C_{12}H_{22}O_{11}**
- Molecular Weight: 342.29648
- ALogP: -4.361
- Rotatable Bonds: 6
- Acceptors: 11
- Donors: 8

**Model Prediction**
- Prediction: Non-Carcinogen
- Probability: 0.216
- Enrichment: 0.673
- Bayesian Score: -2.679

*Prediction*: Positive if the Bayesian score is above the estimated best cutoff value.
*Probability*: The estimated probability that the sample is in the category.
*Enrichment*: An estimate of enrichment, that is, the increased likelihood (versus random) of this sample being in the category.
*Bayesian Score*: The standard Laplacian-modified Bayesian score.

**Mutagenecity**

![C7 Inulin](image)

- **C_{12}H_{22}O_{11}**
- Molecular Weight: 342.29648
- ALogP: -4.361
- Rotatable Bonds: 6
- Acceptors: 11
- Donors: 8

**Model Prediction**
- Prediction: Non-Mutagen
- Probability: 0.322
- Enrichment: 0.577
- Bayesian Score: -11.346

*Prediction*: Positive if the Bayesian score is above the estimated best cutoff value.
*Probability*: The estimated probability that the sample is in the category.
*Enrichment*: An estimate of enrichment, that is, the increased likelihood (versus random) of this sample being in the category.
*Bayesian Score*: The standard Laplacian-modified Bayesian score.
Rat oral LD50 (g/Kg Body weight)  

**Chemical Structure 1**

- Chemical formula: C_{12}H_{22}O_{11}
- Molecular Weight: 342.29648
- ALogP: -4.361
- Rotatable Bonds: 6
- Acceptors: 11
- Donors: 8

**Model Prediction**

- Prediction: 20.789
- Unit: g/kg_body_weight

---

**Chemical Structure 2**

- Chemical formula: C_{12}H_{22}O_{11}
- Molecular Weight: 342.29648
- ALogP: -4.361
- Rotatable Bonds: 6
- Acceptors: 11
- Donors: 8

**Model Prediction**

- Prediction: 0.000
- Unit: g/kg_body_weight
Model Prediction
Prediction: Mild
Probability: 0.159
Enrichment: 0.432
Bayesian Score: -6.949

Prediction: Non-Sensitizer
Probability: 0.670
Enrichment: 0.977
Bayesian Score: -2.394

C_{12}H_{22}O_{11}
Molecular Weight: 342.29648
 ALOGP: -4.361
Rotatable Bonds: 6
Acceptors: 11
Donors: 8

C_{12}H_{22}O_{11}
Molecular Weight: 342.29648
 ALOGP: -4.361
Rotatable Bonds: 6
Acceptors: 11
Donors: 8

Prediction: Positive if the Bayesian score is above the estimated best cutoff value.
Probability: The estimated probability that the sample is in the category.
Enrichment: An estimate of enrichment, that is, the increased likelihood (versus random) of this sample being in the category.
Bayesian Score: The standard Laplacian-modified Bayesian score.
Aerobic Biodegradability

\[ \text{C}_{12}\text{H}_{22}\text{O}_{11} \]
Molecular Weight: 342.29648
\( \text{ALogP}: -4.361 \)
Rotatable Bonds: 6
Acceptors: 11
Donors: 8

**Model Prediction**

*Prediction*: Degradable

*Probability*: 0.609

*Enrichment*: 1.395

*Bayesian Score*: 2.532

Prediction: Positive if the Bayesian score is above the estimated best cutoff value.
Probability: The estimated probability that the sample is in the category.
Enrichment: An estimate of enrichment, that is, the increased likelihood (versus random) of this sample being in the category.
Bayesian Score: The standard Laplacian-modified Bayesian score.
Supplementary figure 3(h) - C8 - Poly Cysteine

FDA Rodent Carcinogenicity

Model Prediction
Prediction: Non-Carcinogen
Probability: 0.210
Enrichment: 0.656
Bayesian Score: -4.229

Prediction: Positive if the Bayesian score is above the estimated best cutoff value.
Probability: The estimated probability that the sample is in the category.
Enrichment: An estimate of enrichment, that is, the increased likelihood (versus random) of this sample being in the category.
Bayesian Score: The standard Laplacian-modified Bayesian score.

Mutagenecity

Model Prediction
Prediction: Non-Mutagen
Probability: 0.631
Enrichment: 1.130
Bayesian Score: -4.187

Prediction: Positive if the Bayesian score is above the estimated best cutoff value.
Probability: The estimated probability that the sample is in the category.
Enrichment: An estimate of enrichment, that is, the increased likelihood (versus random) of this sample being in the category.
Bayesian Score: The standard Laplacian-modified Bayesian score.
Rat oral LD50 (g/Kg Body weight)

C₂H₇NO₂S
Molecular Weight: 121.15818
LogP: -3.078
Rotatable Bonds: 2
Acceptors: 4
Donors: 3

Model Prediction
Prediction: 0.514
Unit: g/kg_body_weight

Rat Maximum tolerated dose (g/Kg Body weight)

C₂H₇NO₂S
Molecular Weight: 121.15818
LogP: -3.078
Rotatable Bonds: 2
Acceptors: 4
Donors: 3

Model Prediction
Prediction: 0.653
Unit: g/kg_body_weight
Skin Irritancy

![Chemical Structure 1](image1)

C₆H₆NO₂S
Molecular Weight: 121.15818
ALogP: -3.078
Rotatable Bonds: 2
Acceptors: 4
Donors: 3

Model Prediction
Prediction: Non-Irritant
Probability: 0.971
Enrichment: 1.055
Bayesian Score: -0.704

Prediction: Positive if the Bayesian score is above the estimated best cutoff value.
Probability: The estimated probability that the sample is in the category.
Enrichment: An estimate of enrichment, that is, the increased likelihood (versus random) of this sample being in the category.
Bayesian Score: The standard Laplacian-modified Bayesian score.

![Chemical Structure 2](image2)

C₆H₆NO₂S
Molecular Weight: 121.15818
ALogP: -3.078
Rotatable Bonds: 2
Acceptors: 4
Donors: 3

Model Prediction
Prediction: Non-Sensitizer
Probability: 0.725
Enrichment: 1.057
Bayesian Score: -1.390

Prediction: Positive if the Bayesian score is above the estimated best cutoff value.
Probability: The estimated probability that the sample is in the category.
Enrichment: An estimate of enrichment, that is, the increased likelihood (versus random) of this sample being in the category.
Bayesian Score: The standard Laplacian-modified Bayesian score.

Skin sensitization
Aerobic Biodegradability

C₆H₈NO₂S
Molecular Weight: 121.16818
ALogP: -3.078
Rotatable Bonds: 2
Acceptors: 4
Donors: 3

Model Prediction
Prediction: Degradable
Probability: 0.562
Enrichment: 1.289
Bayesian Score: 1.548

Prediction: Positive if the Bayesian score is above the estimated best cutoff value.
Probability: The estimated probability that the sample is in the category.
Enrichment: An estimate of enrichment, that is, the increased likelihood (versus random) of this sample being in the category.
Bayesian Score: The standard Laplacian-modified Bayesian score.
Supplementary figure 3(i)- C9 – Chitosan

FDA Rodent Carcinogenicity

<table>
<thead>
<tr>
<th>Model Prediction</th>
<th>Prediction: Non-Carcinogen</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probability:</td>
<td>0.215</td>
</tr>
<tr>
<td>Enrichment:</td>
<td>0.672</td>
</tr>
<tr>
<td>Bayesian Score:</td>
<td>-2.754</td>
</tr>
</tbody>
</table>

Prediction: Positive if the Bayesian score is above the estimated best cutoff value. Probability: The estimated probability that the sample is in the category. Enrichment: An estimate of enrichment, that is, the increased likelihood (versus random) of this sample being in the category. Bayesian Score: The standard Laplacian-modified Bayesian score.

Mutagenecity

<table>
<thead>
<tr>
<th>Model Prediction</th>
<th>Prediction: Non-Mutagen</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probability:</td>
<td>0.711</td>
</tr>
<tr>
<td>Enrichment:</td>
<td>1.273</td>
</tr>
<tr>
<td>Bayesian Score:</td>
<td>-1.342</td>
</tr>
</tbody>
</table>

Prediction: Positive if the Bayesian score is above the estimated best cutoff value. Probability: The estimated probability that the sample is in the category. Enrichment: An estimate of enrichment, that is, the increased likelihood (versus random) of this sample being in the category. Bayesian Score: The standard Laplacian-modified Bayesian score.
Rat oral LD50 (g/Kg Body weight)  
Rat Maximum tolerated dose (g/Kg Body weight)

\[
\text{C}_6\text{H}_{13}\text{NO}_6 \\
\text{Molecular Weight: 179.17112} \\
\text{ALogP: -2.804} \\
\text{Rotatable Bonds: 1} \\
\text{Acceptors: 6} \\
\text{Donors: 5}
\]

Model Prediction  
Prediction: 3.241  
Unit: g/kg_body_weight

\[
\text{C}_6\text{H}_{13}\text{NO}_6 \\
\text{Molecular Weight: 179.17112} \\
\text{ALogP: -2.804} \\
\text{Rotatable Bonds: 1} \\
\text{Acceptors: 6} \\
\text{Donors: 5}
\]

Model Prediction  
Prediction: 0.268  
Unit: g/kg_body_weight
Model Prediction

**Prediction: Mild**
Probability: 0.242
Enrichment: 0.659
Bayesian Score: -4.415

Predicted: Positive if the Bayesian score is above the estimated best cutoff value.
Probability: The estimated probability that the sample is in the category.
Enrichment: An estimate of enrichment, that is, the increased likelihood (versus random) of this sample being in the category.
Bayesian Score: The standard Laplacian-modified Bayesian score.

Model Prediction

**Prediction: Non-Sensitizer**
Probability: 0.631
Enrichment: 0.920
Bayesian Score: -3.018

Predicted: Positive if the Bayesian score is above the estimated best cutoff value.
Probability: The estimated probability that the sample is in the category.
Enrichment: An estimate of enrichment, that is, the increased likelihood (versus random) of this sample being in the category.
Bayesian Score: The standard Laplacian-modified Bayesian score.
Aerobic Biodegradability

C₈H₁₂NO₅
Molecular Weight: 179.17112
ALogP: -2.804
Rotatable Bonds: 1
Acceptors: 6
Donors: 5

Model Prediction
Prediction: Degradable
Probability: 0.620
Enrichment: 1.421
Bayesian Score: 2.775

Prediction: Positive if the Bayesian score is above the estimated best cutoff value.
Probability: The estimated probability that the sample is in the category.
Enrichment: An estimate of enrichment, that is, the increased likelihood (versus random) of this sample being in the category.
Bayesian Score: The standard Laplacian-modified Bayesian score.
Supplementary figure 3(j) - C10 - Pectin

**FDA Rodent Carcinogenicity**

![Molecular Structure 1](image1)

\[ C_6H_{10}O_7 \]
Molecular Weight: 194.1394
ALogP: -2.386
Rotatable Bonds: 1
Acceptors: 7
Donors: 5

**Model Prediction**

**Prediction:** Non-Carcinogen
Probability: 0.213
Enrichment: 0.666
Bayesian Score: -3.131

*Prediction:* Positive if the Bayesian score is above the estimated best cutoff value.
*Probability:* The estimated probability that the sample is in the category.
*Enrichment:* An estimate of enrichment, that is, the increased likelihood (versus random) of this sample being in the category.
*Bayesian Score:* The standard Laplacian-modified Bayesian score.

**FDA Rodent Mutagenecity**

![Molecular Structure 2](image2)

\[ C_6H_{10}O_7 \]
Molecular Weight: 194.1394
ALogP: -2.386
Rotatable Bonds: 1
Acceptors: 7
Donors: 5

**Model Prediction**

**Prediction:** Non-Mutagen
Probability: 0.567
Enrichment: 1.016
Bayesian Score: -5.935

*Prediction:* Positive if the Bayesian score is above the estimated best cutoff value.
*Probability:* The estimated probability that the sample is in the category.
*Enrichment:* An estimate of enrichment, that is, the increased likelihood (versus random) of this sample being in the category.
*Bayesian Score:* The standard Laplacian-modified Bayesian score.
Rat oral LD50 (g/Kg Body weight)

\[
\text{C}_6\text{H}_{10}\text{O}_7
\]
Molecular Weight: 194.1394
\(\text{ALogP: } -2.386\)
Rotatable Bonds: 1
Acceptors: 7
Donors: 5

Model Prediction
Prediction: 0.525
Unit: g/kg_body_weight

Rat Maximum tolerated dose (g/Kg Body weight)

\[
\text{C}_6\text{H}_{10}\text{O}_7
\]
Molecular Weight: 194.1384
\(\text{ALogP: } -2.386\)
Rotatable Bonds: 1
Acceptors: 7
Donors: 5

Model Prediction
Prediction: 3.576
Unit: g/kg_body_weight
Skin Irritancy

\[ \text{C}_8\text{H}_{16}\text{O}_7 \]
Molecular Weight: 194.1394
ALogP: 2.386
Rotatable Bonds: 1
Acceptors: 7
Donors: 5

Model Prediction
Prediction: Non-Irritant
Probability: 0.971
Enrichment: 1.055
Bayesian Score: -0.726

Prediction: Positive if the Bayesian score is above the estimated best cutoff value.
Probability: The estimated probability that the sample is in the category.
Enrichment: An estimate of enrichment, that is, the increased likelihood (versus random) of this sample being in the category.
Bayesian Score: The standard Laplacian-modified Bayesian score.

Skin sensitization

\[ \text{C}_8\text{H}_{16}\text{O}_7 \]
Molecular Weight: 194.1394
ALogP: 2.386
Rotatable Bonds: 1
Acceptors: 7
Donors: 5

Model Prediction
Prediction: Non-Sensitizer
Probability: 0.704
Enrichment: 1.026
Bayesian Score: -1.792

Prediction: Positive if the Bayesian score is above the estimated best cutoff value.
Probability: The estimated probability that the sample is in the category.
Enrichment: An estimate of enrichment, that is, the increased likelihood (versus random) of this sample being in the category.
Bayesian Score: The standard Laplacian-modified Bayesian score.
Aerobic Biodegradability

C₆H₁₀O₇
Molecular Weight: 194.1394
 ALOGP: -2.386
 Rotatable Bonds: 1
 Acceptors: 7
 Donors: 5

Model Prediction
Prediction: Degradable
Probability: 0.632
Enrichment: 1.447
Bayesian Score: 3.026

Prediction: Positive if the Bayesian score is above the estimated best cutoff value.
Probability: The estimated probability that the sample is in the category.
Enrichment: An estimate of enrichment, that is, the increased likelihood (versus random) of this sample being in the category.
Bayesian Score: The standard Laplacian-modified Bayesian score.
**Supplementary figure 3(k) - C11 - Poly(propylene glycol)**

**FDA Rodent Carcinogenicity**

- **Formula:** C₈H₁₄O₃
- **Molecular Weight:** 134.17356
- **ALogP:** -0.274
- **Rotatable Bonds:** 4
- **Acceptors:** 3
- **Donors:** 2

**Model Prediction**
- **Prediction:** Non-Carcinogen
- **Probability:** 0.222
- **Enrichment:** 0.692
- **Bayesian Score:** -1.865

**Definition:**
- **Prediction:** Positive if the Bayesian score is above the estimated best cutoff value.
- **Probability:** The estimated probability that the sample is in the category.
- **Enrichment:** An estimate of enrichment, that is, the increased likelihood (versus random) of this sample being in the category.
- **Bayesian Score:** The standard Laplacian-modified Bayesian score.

**Mutagenecity**

- **Formula:** C₈H₁₄O₃
- **Molecular Weight:** 134.17356
- **ALogP:** -0.274
- **Rotatable Bonds:** 4
- **Acceptors:** 3
- **Donors:** 2

**Model Prediction**
- **Prediction:** Non-Mutagen
- **Probability:** 0.686
- **Enrichment:** 1.229
- **Bayesian Score:** -2.332

**Definition:**
- **Prediction:** Positive if the Bayesian score is above the estimated best cutoff value.
- **Probability:** The estimated probability that the sample is in the category.
- **Enrichment:** An estimate of enrichment, that is, the increased likelihood (versus random) of this sample being in the category.
- **Bayesian Score:** The standard Laplacian-modified Bayesian score.
Rat oral LD50 (g/Kg Body weight)  

![Chemical Structure 1][1]

\[ C_6H_{14}O_3 \]
Molecular Weight: 134.17356
ALogP: -0.274
Rotatable Bonds: 4
Acceptors: 3
Donors: 2

**Model Prediction**
Prediction: 12.098
Unit: g/kg_body_weight

Rat Maximum tolerated dose (g/Kg Body weight)

![Chemical Structure 2][2]

\[ C_6H_{14}O_3 \]
Molecular Weight: 134.17356
ALogP: -0.274
Rotatable Bonds: 4
Acceptors: 3
Donors: 2

**Model Prediction**
Prediction: 0.187
Unit: g/kg_body_weight
**Model Prediction**

**Prediction:** Mild  
**Probability:** 0.244  
**Enrichment:** 0.664  
**Bayesian Score:** -4.358

**Prediction:** Non-Sensitizer  
**Probability:** 0.513  
**Enrichment:** 0.747  
**Bayesian Score:** -4.681

*Prediction:* Positive if the Bayesian score is above the estimated best cutoff value.  
*Probability:* The estimated probability that the sample is in the category.  
*Enrichment:* An estimate of enrichment, that is, the increased likelihood (versus random) of this sample being in the category.  
*Bayesian Score:* The standard Laplacian-modified Bayesian score.
Aerobic Biodegradability

C₆H₁₄O₃
Molecular Weight: 134.17356
ALogP: -0.274
Rotatable Bonds: 4
Acceptors: 3
Donors: 2

**Model Prediction**

**Prediction:** Degradable
**Probability:** 0.711
**Enrichment:** 1.629
**Bayesian Score:** 4.861

*Prediction:* Positive if the Bayesian score is above the estimated best cutoff value.
*Probability:* The estimated probability that the sample is in the category.
*Enrichment:* An estimate of enrichment, that is, the increased likelihood (versus random) of this sample being in the category.
*Bayesian Score:* The standard Laplacian-modified Bayesian score.
**Supplementary figure 3(l) - C12 - Poly(propylene imine)**

**FDA Rodent Carcinogenicity**

- 

**Mutagenecity**

- 

**Model Prediction**

**Prediction: Non-Carcinogen**
- Probability: 0.558
- Enrichment: 1.417
- Bayesian Score: -0.647

**Prediction: Non-Mutagen**
- Probability: 0.666
- Enrichment: 1.192
- Bayesian Score: -3.064
Rat oral LD50 (g/Kg Body weight)  

<table>
<thead>
<tr>
<th>Chemical Structure</th>
<th>Molecular Weight</th>
<th>ALogP</th>
<th>Rotatable Bonds</th>
<th>Acceptors</th>
<th>Donors</th>
</tr>
</thead>
<tbody>
<tr>
<td>C₃H₇⁺N</td>
<td>57.09438</td>
<td>0.131</td>
<td>3</td>
<td>3</td>
<td>1</td>
</tr>
</tbody>
</table>

Model Prediction  
Prediction: 0.055  
Unit: g/kg_body_weight

Rat Maximum tolerated dose (g/Kg Body weight)  

<table>
<thead>
<tr>
<th>Chemical Structure</th>
<th>Molecular Weight</th>
<th>ALogP</th>
<th>Rotatable Bonds</th>
<th>Acceptors</th>
<th>Donors</th>
</tr>
</thead>
<tbody>
<tr>
<td>C₃H₇⁺⁺N</td>
<td>57.09438</td>
<td>0.131</td>
<td>3</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Model Prediction  
Prediction: 0.069  
Unit: g/kg_body_weight
<table>
<thead>
<tr>
<th>Model Prediction</th>
<th>Model Prediction</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Prediction:</strong></td>
<td><strong>Prediction:</strong></td>
</tr>
<tr>
<td>Mild</td>
<td>Non-Sensitizer</td>
</tr>
<tr>
<td><strong>Probability:</strong></td>
<td><strong>Probability:</strong></td>
</tr>
<tr>
<td>0.275</td>
<td>0.681</td>
</tr>
<tr>
<td><strong>Enrichment:</strong></td>
<td><strong>Enrichment:</strong></td>
</tr>
<tr>
<td>0.748</td>
<td>0.993</td>
</tr>
<tr>
<td><strong>Bayesian Score:</strong></td>
<td><strong>Bayesian Score:</strong></td>
</tr>
<tr>
<td>-3.519</td>
<td>-2.202</td>
</tr>
</tbody>
</table>

Prediction: Predict if the Bayesian score is above the estimated best cutoff value.
Probability: The estimated probability that the sample is in the category.
Enrichment: An estimate of enrichment, that is, the increased likelihood (versus random) of this sample being in the category.
Bayesian Score: The standard Laplacian-modified Bayesian score.
C₃H₅N₂
Molecular Weight: 57.09438
ALogP: 0.131
Rotatable Bonds: 3
Acceptors: 1
Donors: 1

**Model Prediction**

*Prediction: Non-Degradable*

Probability: 0.514

Enrichment: 1.177

Bayesian Score: 0.529

Prediction: Positive if the Bayesian score is above the estimated best cutoff value.
Probability: The estimated probability that the sample is in the category.
Enrichment: An estimate of enrichment, that is, the increased likelihood (versus random) of this sample being in the category.
Bayesian Score: The standard Laplacian-modified Bayesian score.
Supplementary figure 3(m): C13 - Poly (lactic-co-glycolic acid)

**FDA Rodent Carcinogenicity**

- **Formula:** C₈H₈O₅
- **Molecular Weight:** 148.11402
- **ALogP:** -0.686
- **Rotatable Bonds:** 4
- **Acceptors:** 5
- **Donors:** 2

**Model Prediction**

**Prediction:** Non-Carcinogen

**Probability:** 0.213

**Enrichment:** 0.664

**Bayesian Score:** -3.290

_**Prediction:** Positive if the Bayesian score is above the estimated best cutoff value._

_**Probability:** The estimated probability that the sample is in the category._

_**Enrichment:** An estimate of enrichment, that is, the increased likelihood (versus random) of this sample being in the category._

_**Bayesian Score:** The standard Laplacian-modified Bayesian score._

---

**Mutagenecity**

- **Formula:** C₈H₈O₅
- **Molecular Weight:** 148.11402
- **ALogP:** -0.686
- **Rotatable Bonds:** 4
- **Acceptors:** 5
- **Donors:** 2

**Model Prediction**

**Prediction:** Non-Mutagen

**Probability:** 0.543

**Enrichment:** 0.972

**Bayesian Score:** -6.541

_**Prediction:** Positive if the Bayesian score is above the estimated best cutoff value._

_**Probability:** The estimated probability that the sample is in the category._

_**Enrichment:** An estimate of enrichment, that is, the increased likelihood (versus random) of this sample being in the category._

_**Bayesian Score:** The standard Laplacian-modified Bayesian score._
Rat oral LD50 (g/Kg Body weight)

C₂H₆O₅
Molecular Weight: 148.11402
ALogP: -0.686
Rotatable Bonds: 4
Acceptors: 5
Donors: 2

Model Prediction
Prediction: 2.992
Unit: g/kg_body_weight

Rat Maximum tolerated dose (g/Kg Body weight)

C₂H₆O₅
Molecular Weight: 148.11402
ALogP: -0.686
Rotatable Bonds: 4
Acceptors: 5
Donors: 2

Model Prediction
Prediction: 0.427
Unit: g/kg_body_weight
Skin Irritancy

C₂H₂O₅
Molecular Weight: 148.11402
ALogP: -0.686
Rotatable Bonds: 4
Acceptors: 5
Donors: 2

Model Prediction
Prediction: Non-Irritant
Probability: 0.972
Enrichment: 1.056
Bayesian Score: -0.615

Prediction: Positive if the Bayesian score is above the estimated best cutoff value.
Probability: The estimated probability that the sample is in the category.
Enrichment: An estimate of enrichment, that is, the increased likelihood (versus random) of this sample being in the category.
Bayesian Score: The standard Laplacian-modified Bayesian score.

C₂H₂O₅
Molecular Weight: 148.11402
ALogP: -0.686
Rotatable Bonds: 4
Acceptors: 5
Donors: 2

Model Prediction
Prediction: Non-Sensitizer
Probability: 0.716
Enrichment: 1.044
Bayesian Score: -1.565

Prediction: Positive if the Bayesian score is above the estimated best cutoff value.
Probability: The estimated probability that the sample is in the category.
Enrichment: An estimate of enrichment, that is, the increased likelihood (versus random) of this sample being in the category.
Bayesian Score: The standard Laplacian-modified Bayesian score.
Aerobic Biodegradability

\[
\text{C}_{2}\text{H}_{5}\text{O}_5
\]

Molecular Weight: 148.11402
\text{ALogP}: -0.666
Rotatable Bonds: 4
Acceptors: 5
Donors: 2

**Model Prediction**

Prediction: Degradable
Probability: 0.717
Enrichment: 1.644
Bayesian Score: 5.017

Prediction: Positive if the Bayesian score is above the estimated best cutoff value.
Probability: The estimated probability that the sample is in the category.
Enrichment: An estimate of enrichment, that is, the increased likelihood (versus random) of this sample being in the category.
Bayesian Score: The standard Laplacian-modified Bayesian score.
Supplementary figure 3(n)- C14 - Deoxycholic acid

FDA Rodent Carcinogenicity

Mutagenecity

\[
\text{C}_{24}\text{H}_{40}\text{O}_{4}
\]
Molecular Weight: 392.57199
ALogP: 4.082
Rotatable Bonds: 4
Acceptors: 4
Donors: 3

Model Prediction
Prediction: Non-Carcinogen
Probability: 0.227
Enrichment: 0.707
Bayesian Score: -1.366

Prediction: Positive if the Bayesian score is above the estimated best cutoff value.
Probability: The estimated probability that the sample is in the category.
Enrichment: An estimate of enrichment, that is, the increased likelihood (versus random) of this sample being in the category.
Bayesian Score: The standard Laplacian-modified Bayesian score.

Model Prediction
Prediction: Non-Mutagen
Probability: 0.000
Enrichment: 0.000
Bayesian Score: -64.641

Prediction: Positive if the Bayesian score is above the estimated best cutoff value.
Probability: The estimated probability that the sample is in the category.
Enrichment: An estimate of enrichment, that is, the increased likelihood (versus random) of this sample being in the category.
Bayesian Score: The standard Laplacian-modified Bayesian score.
Rat oral LD50 (g/Kg Body weight)  

Rat Maximum tolerated dose (g/Kg Body weight)

C_{24}H_{42}O_{4}
Molecular Weight: 392.57199
ALogP: 4.082
Rotatable Bonds: 4
Acceptors: 4
Donors: 3

Model Prediction
Prediction: 6.358
Unit: g/kg_body_weight

C_{24}H_{42}O_{4}
Molecular Weight: 392.57199
ALogP: 4.082
Rotatable Bonds: 4
Acceptors: 4
Donors: 3

Model Prediction
Prediction: 0.190
Unit: g/kg_body_weight
**Model Prediction**

**Skin Irritancy**

- **Prediction**: Moderate - Severe
- **Probability**: 0.472
- **Enrichment**: 1.283
- **Bayesian Score**: 1.758

- Prediction: Positive if the Bayesian score is above the estimated best cutoff value.
- Probability: The estimated probability that the sample is in the category.
- Enrichment: An estimate of enrichment, that is, the increased likelihood (versus random) of this sample being in the category.
- Bayesian Score: The standard Laplacian-modified Bayesian score.

**Skin Sensitization**

- **Prediction**: Weak - Sensitizer
- **Probability**: 0.185
- **Enrichment**: 0.238
- **Bayesian Score**: -8.518

- Prediction: Positive if the Bayesian score is above the estimated best cutoff value.
- Probability: The estimated probability that the sample is in the category.
- Enrichment: An estimate of enrichment, that is, the increased likelihood (versus random) of this sample being in the category.
- Bayesian Score: The standard Laplacian-modified Bayesian score.
Aerobic Biodegradability

C_{24}H_{40}O_4
Molecular Weight: 392.57199
ALogP: 4.082
Rotatable Bonds: 4
Acceptors: 4
Donors: 3

Model Prediction
Prediction: Degradable
Probability: 0.837
Enrichment: 1.919
Bayesian Score: 8.551

Prediction: Positive if the Bayesian score is above the estimated best cutoff value.
Probability: The estimated probability that the sample is in the category.
Enrichment: An estimate of enrichment, that is, the increased likelihood (versus random) of this sample being in the category.
Bayesian Score: The standard Laplacian-modified Bayesian score.
Supplementary figure 4
Toxicity studies for Drug delivering molecules by Toxtree: C1 - Vitamin B12 (cobalamin); C2 - Vitamin H (Biotin); C3 - Folic acid (Vitamin M / Vitamin B9); C4 - Vitamin B1 (Thiamin); C5 - L-Carnitine (Vitamin B7); C6 - Poly-N-vinylpyrrolidone; C7 - Inulin; C8 - Poly Cysteine; C9 - Chitosan; C10 - Pectin; C11 - Poly(propylene glycol); C12 - Poly(propylene imine); C13 - Poly (lactic-co-glycolic acid); C14 - Deoxycholic acid

Supplementary figure 4(a) - C1 - Vitamin B12 (cobalamin)

Supplementary figure 4(b) - C2 - Vitamin H (Biotin)
Supplementary figure 4(c) - C3 - Folic acid (Vitamin M / Vitamin B9)

Supplementary figure 4(d) - C4 - Vitamin B1 (Thiamin)
Supplementary figure 4(e) - C5 - L-Carnitine (Vitamin BT)

Supplementary figure 4(f) - C6 - Poly-N-vinylpyrrolidone
Supplementary figure 4(g) - C7; Inulin

Supplementary figure 3(h) - C8 - Poly Cysteine
Supplementary figure 4(i)- C9 – Chitosan

Supplementary figure 4(j)- C10 – Pectin
Supplementary figure 4(k) - C11 - Poly(propylene glycol)

Supplementary figure 4(l) - C12 - Poly(propylene imine)
Supplementary figure 4(m) - C13 - Poly (lactic-co-glycolic acid)

Supplementary figure 4(n) - C14 - Deoxycholic acid
Supplementary figure 5
Conjugation results of Human Insulin Monomer (PDB ID: 1HLS), with all listed drug delivering molecules individually by Discovery Studio software. C1 - Vitamin B12 (cobalamin); C2 - Vitamin H (Biotin); C3 - Folic acid (Vitamin M / Vitamin B9); C4 - Vitamin B1 (Thiamin); C5 - L-Carnitine (Vitamin BT); C6 - Poly-N-vinylpyrrolidone; C7 - Inulin; C8 - Poly Cysteine; C9 - Chitosan; C10 - Pectin; C11 - Poly(propylene glycol); C12 - Poly(propylene imine); C13 - Poly (lactic-co-glycolic acid); C14 - Deoxycholic acid

Supplementary figure 5(a)- C1 - Vitamin B12 (cobalamin)
No Conjugation

Supplementary figure 5(b)- C2 - Vitamin H (Biotin)

Supplementary figure 5(c)- C3 - Folic acid (Vitamin M / Vitamin B9)
No Conjugation
Supplementary figure 5(d) - C4 - Vitamin B1 (Thiamin)

Supplementary figure 5(e) - C5 - L-Carnitine (Vitamin BT)

Supplementary figure 5(f) - C6 - Poly-N-vinylpyrrolidone

No Conjugation
**Supplementary figure 5(g): C7 - Inulin**

**Supplementary figure 5(h): C8 - Poly Cysteine**
Supplementary figure 5(i) - C9 – Chitosan

Supplementary figure 5(j) - C10 – Pectin
Supplementary figure 5(k)- C11 - Poly(propylene glycol)

Supplementary figure 5(l)- C12 - Poly(propylene imine)
No Conjugation

Supplementary figure 5(m)- C13 - Poly(lactic-co-glycolic acid)
No Conjugation

Supplementary figure 5(n)- C14 - Deoxycholic acid
No Conjugation
Supplementary figure 6
Conjugation results of Human insulin hexamer (PDB ID: 1AIO), with all listed drug delivering molecules individually by Discovery Studio software. C1 - Vitamin B12 (cobalamin); C2 - Vitamin H (Biotin); C3 - Folic acid (Vitamin M / Vitamin B9); C4 - Vitamin B1 (Thiamin); C5 - L-Carnitine (Vitamin BT); C6 - Poly-N-vinylpyrrolidone; C7 - Inulin; C8 - Poly Cysteine; C9 - Chitosan; C10 - Pectin; C11 - Poly(propylene glycol); C12 - Poly(propylene imine); C13 - Poly (lactic-co-glycolic acid); C14 - Deoxycholic acid

Supplementary figure 6(a)- C1 - Vitamin B12 (cobalamin)
No Conjugation

Supplementary figure 6(b)- C2 - Vitamin H (Biotin)
Supplementary figure 6(c) - C3 - Folic acid (Vitamin M / Vitamin B9)

Supplementary figure 6(d) - C4 - Vitamin B1 (Thiamin)
Supplementary figure 6(e) - C5 - L-Carnitine (Vitamin BT)

Supplementary figure 6(f) - C6 - Poly-N-vinylpyrrolidone
No Conjugation

Supplementary figure 6(g) - C7; Inulin
**Supplementary figure 6(j)- C10 - Pectin**

**Supplementary figure 6(k)- C11 - Poly(propylene glycol)**

**Supplementary figure 6(l)- C12 - Poly(propylene imine)**

No Conjugation
Supplementary figure 6(m) - C13 - Poly (lactic-co-glycolic acid)

Supplementary figure 6(n) - C14 - Deoxycholic acid
Supplementary figure 7
Conjugation results of Insulin Lispro (PDB ID: 1LPH), with all listed drug delivering molecules individually by Discovery Studio software. C1 - Vitamin B12 (cobalamin); C2 - Vitamin H (Biotin); C3 - Folic acid (Vitamin M / Vitamin B9); C4 - Vitamin B1 (Thiamin); C5 - L-Carnitine (Vitamin BT); C6 - Poly-N-vinylpyrrolidone; C7 - Inulin; C8 - Poly Cysteine; C9 - Chitosan; C10 - Pectin; C11 - Poly(propylene glycol); C12 - Poly(propylene imine); C13 - Poly (lactic-co-glycolic acid); C14 - Deoxycholic acid

Supplementary figure 7(a) - C1 - Vitamin B12 (cobalamin)
No Conjugation

Supplementary figure 7(b) - C2 - Vitamin H (Biotin)
No Conjugation

Supplementary figure 7(c) - C3 - Folic acid (Vitamin M / Vitamin B9)
Supplementary figure 7(d) - C4 - Vitamin B1 (Thiamin)

Supplementary figure 7(e) - C5 - L-Carnitine (Vitamin BT)

Supplementary figure 7(f) - C6 - Poly-N-vinylpyrrolidone
No Conjugation
Supplementary figure 7(i) - C9 – Chitosan

Supplementary figure 7(ii) - C10 – Pectin
Supplementary figure 7(k)- C11 - Poly(propylene glycol)

Supplementary figure 7(l)- C12 - Poly(propylene imine)
No Conjugation

Supplementary figure 7(m)- C13 - Poly (lactic-co-glycolic acid)

Supplementary figure 7(n)- C14 - Deoxycholic acid
No Conjugation
Supplementary figure 8

Supplementary figure 8(a)
Interaction results of Oral insulin conjugates (Insulin Monomer (1HLS)- DDM Conjugates) with Insulin Receptor (IR). It does not show any interaction in leucine-rich repeat domain (L1, residues 1-157) and in C-terminus of the α-chain (αCT, residues 704-715).
Supplementary figure 8(b)
Interaction results of Oral insulin conjugates (Insulin Hexamer (1AIO)- DDM Conjugates)- DDM Conjugates with Insulin Receptor (IR). It shows the interaction in ARG86, ASN34 of leucine-rich repeat domain (L1, residues 1-157) and no interaction in C-terminus of the α-chain (αCT, residues 704-715).
Supplementary figure 8(c)

Interaction results of Oral insulin conjugates (Insulin Lispro (1LPH) - DDM Conjugates) with Insulin Receptor (IR). It shows the interaction in ARG86, ASN90 and ARG114 of leucine-rich repeat domain (L1, residues 1-157) and no interaction in C-terminus of the α-chain (αCT, residues 704-715).