Drug Discovery

RESEARCH

ISSN 2278 – 540X EISSN 2278 – 5396

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 R[1,*]

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 Pharmacophoric 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 monoconjugate-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.[1] 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.[2,5,6] 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 85years ago.[2,3] Multiple doses of Insulin subcutaneous injections on the daily basis making patients painful and create 'fear of pain' among them.[4] 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.[3,4] 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.[4] As of 2014, no products appeared to be successful in the market because of lack of physical stability against biological barriers.[4–11] 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.[12-14] 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 forces[15-18] 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.[8,9] 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.[15] 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.[8] From the earlier studies, we choose that Humulin, monomeric & hexameric 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 Screening[22] of DDAs, PubChem Compound,[23] Zinc Database,[24] KEGG[25] ((Kyoto Encyclopedia of Genes and Genomes)), DrugBank,[26] ChemSpider,[27] ChEMBL,[28] and ChEBI[29] (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-likeness 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)).[32] 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.[33] 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,[34] Vitamin H,[45] Folic acid,[51] Poly-N-vinylpyrrolidone,[67,68] Inulin,[76] Poly Cysteine,[88] Chitosan,[96-98]

Pectin,[105] Poly (Propylene glycol),[112,113] Poly (Propylene imine),[128,129] Poly (lactic-co-glycolic acid),[133,134] Deoxycholic acid[141] 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 (<2000Daltons). 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.[36, 48,53,57,64,69,77,90,97,106,114,130,135,142]

Most of the carriers are having efficient oral bioavailability and intestinal permeability such as Vitamin B12,[36-39] Vitamin H,[46--48] Folic acid,[54,55] Vitamin B1,[58] L-Carnitine,[61-65] Inulin,[78] Poly(propylene glycol),[115,116] Poly(propylene imine),[131] Poly (lactic-co-glycolic acid)[138,139] and Deoxycholic acid.[144,145] Chitosan,[99,100,102,104] and Pectin[108] are graded as moderately efficient in bioavailability while Poly-N-vinylpyrrolidone[70,71] and Poly Cysteine[92,93] 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.[40,48,54,58,65,73,81,94,99,111,117,130,144,145] Among the retrieved carriers, Vitamin B12,[42] Vitamin H,[49] Vitamin B1,[59] Poly-N-vinylpyrrolidone,[74] Inulin,[82,83] Chitosan,[99] Poly(propylene glycol),[122] Poly (lactic-co-glycolic acid)[137] and Deoxycholic acid[146] are chemically inert and they does 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,[44] Vitamin H,[50] Folic acid,[56] Vitamin B1,[60] L-Carnitine,[66] Inulin,[84] Poly Cysteine,[95] Chitosan,[99] Pectin,[109] Poly (lactic-co-glycolic acid)[140] are non-toxic materials and does not produce any untoward reactivity during drug-delivery mechanism. But, Poly-N-vinylpyrrolidone,[75] Poly(propylene glycol),[126] Poly(propylene imine),[132] and Deoxycholic acid[146] 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.[84-87] Generally Inulin[79-81] and Chitosan[99-104] 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.[60] 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.[42] Vitamin H is inert, but it may inert the biological activity of the drug.[49] Poly-N-vinylpyrrolidone,[75] and Poly (propylene imine)[132] are participating in nanoparticle-based drug delivery, and both are moderately toxic. Poly Cysteine is hydrophilic in nature; shows poor permeability across intestinal epithelium;[92,93] 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.[118-122] It may induce Cardiotoxic effects include arrhythmias and cardiac arrest, CNS depression, Renal and hepatic damage has also reported.[123] 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.[124-125] In drug delivering mechanism, it eliminates toxic degradation.[126] It is less toxic than the parent substance (Poly ethylene glycol).[127] Deoxycholic acid is a prominent carrier, which is not inert and moderately toxic; may promote colon tumorigenesis in both animals and humans.[146]

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.[30] 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[79-81] like Chitosan.[99-104] 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.[36-39,78,144,145] 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.

3.1.4. 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 (g/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: 1 LPH) were retrieved from Protein data bank and conjugated with all carriers individually. Nikhil J Kavimandan et al. (2006)[9] and Hinds et al. (2000)[8] 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 competently 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: 1 LPH) 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 B28 & Glu A4 of Insulin Lispro. Inulin was tetra-conjugated fairly with B28Lys, 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 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;[76,78-81] it is non-carcinogen, non-mutagen, low toxic[84] 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;[115,116,122] 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.[124-125] 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[3,8]. 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 Inulin 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

1. Henderson R, Cockbrun I. Scale and Scope in Drug Development: Unpacking the Advantages of Size in Pharmaceutical Research. Journal of Health Economics. 2001; 20(6):1033–1057.
2. Gowthamarajan K, Kulkarni GT. Oral Insulin – Fact or Fiction? Possibilities of achieving oral delivery for insulin. Resonance. 2003;8(5):38-46.
3. Dave N, Hazra P, Khedkar A, Manjunath HS, Iyer H, Suryanarayanan S. Process and purification for manufacture of a modified insulin intended for oral delivery. J Chromatogr A. 2008;1177(2):282-6.
4. Clement S, Still JG, Kosutic G, McAllister RG. Oral insulin product hexyl-insulin monoconjugate 2 (HIM2) in type 1 diabetes mellitus: the glucose stabilization effects of HIM2. Diabetes Technol Ther. 2002;4(4):459-66.
5. Wearley LL. Recent progress in protein and peptide delivery by noninvasive routes. Crit. Rev. Ther. Drug Carrier Syst. 1991;8(4):331-394.

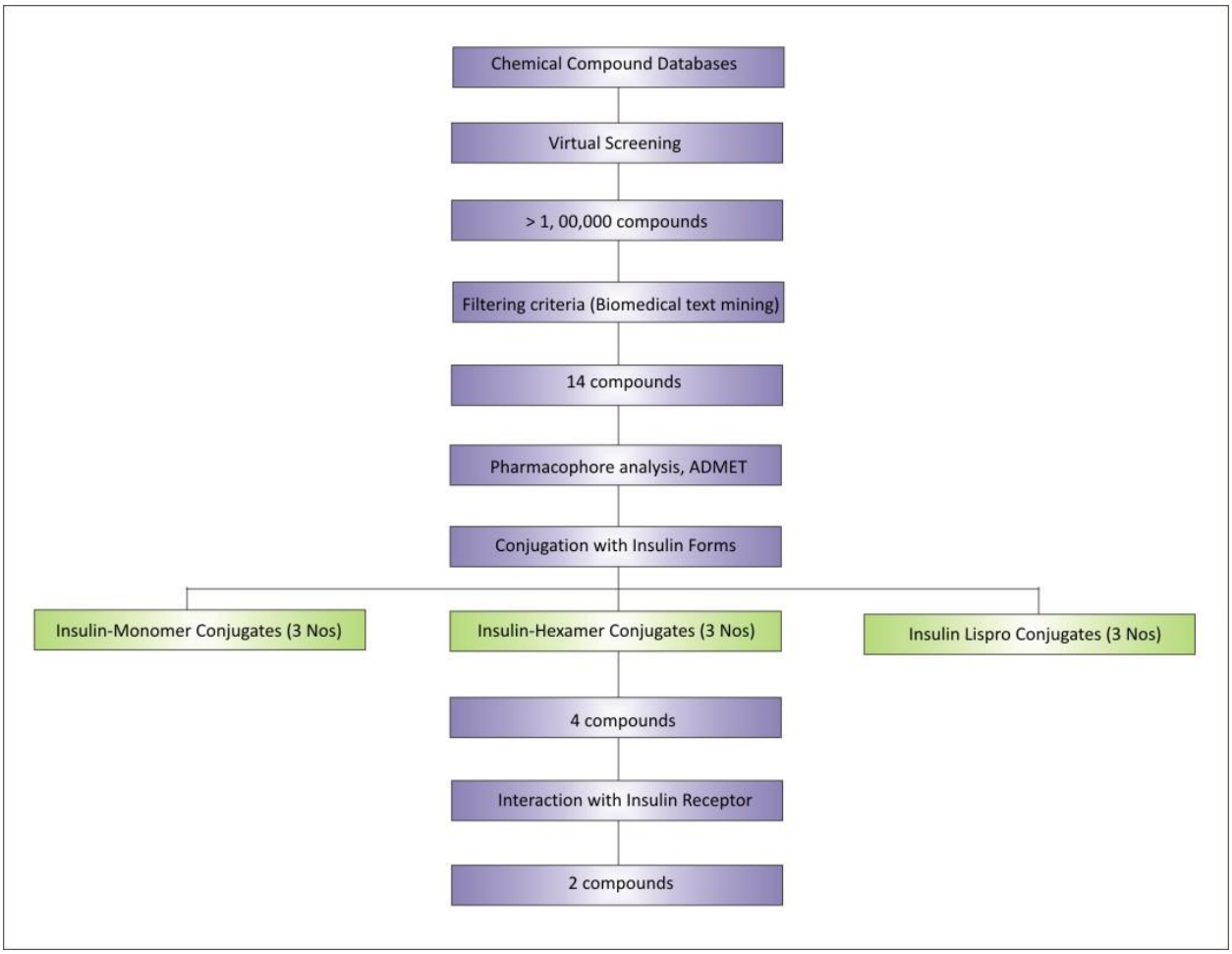
6. Banga AK, Chien YW. Systemic delivery of therapeutic peptides and proteins. Int. J. Pharmacol. 1988;48(1-3):15-50.
7. Lee S, Kim K, Kumar TS, Lee J, Kim SK, Lee DY, Lee YK, Byun Y. Synthesis and biological properties of insulin-deoxycholic acid chemical conjugates. Bioconjug Chem. 2005;16(3):615-20.
8. Hinds K, Koh JJ, Joss L, Liu F, Baudys M, Kim SW. Synthesis and characterization of poly(ethylene glycol)-insulin conjugates. Bioconjug Chem. 2000;11(2):195-201.
9. Kavimandan NJ, Losi E, Wilson JJ, Brodbelt JS, Peppas NA. Synthesis and characterization of insulin-transferrin conjugates. Bioconjug Chem. 2006;17(6):1376-84.
10. Jederstrom G, Andersson A, Gråsjö J, Sjöholm I. Formulating insulin for oral administration: preparation of hyaluronan-insulin complex. Pharm Res. 2004;21(11):2040-7.
11. Anna Lewcoc. Vitamin B12 the key to oral insulin?. 2007 October 23. Available from: http://www-che.syr.edu/faculty/doyle/inpharma.pdf.
12. Gliemann J, Gammeltoft S. The biological activity and the binding affinity of modified insulins determined on isolated rat fat cells. Diabetologia. 1974;10(2), 105-113.
13. Hashimoto M, Takada K, Kiso Y, Muranishi S. Synthesis of palmitoyl derivatives of insulin and their biological activities. Pharm. Res. 1989;6(2), 171-176.
14. Murray-Rust J, McLeod AN, Blundell TL, Wood SP. Structure and evolution of insulins: implications for receptor binding. BioEssays. 1992;14(5), 325-331.
15. Kang S, Brange J, Burch A, Volund A, Owens DR. Subcutaneous insulin absorption explained by insulin's physicochemical properties. Evidence from absorption studies of soluble human insulin and insulin analogues in humans. Diabetes Care. 1991;14(11):942-948.
16. Blundell T, Dodson G, Hodgkin D, Mercola D. Insulin: the structure in the crystal and its reflection in chemistry and biology. Adv. Protein Chem. 1972;26:279-402.
17. Derwenda U, Derwenda Z, Dodson GG, Hubbard RE, Korber F. Molecular structure of insulin: the insulin monomer and its assembly. Br. Med. Bull. 1989;45:4-18.
18. Mark AE, Jeffrey PD. The self-association of zinc-free bovine insulin. Biol. Chem. Hoppe-Seyler. 1990;371:1165-1174.
19. Uchio T, Baudys M, Liu F, Kim SW. Novel site-specific insulin conjugates with enhanced stability and extended action profile for parenteral delivery. Adv. Drug Delivery Rev. 1999;35:289-306.
20. Gerich JE. Novel Insulins: Expanding Options in Diabetes Management. The Merican Journal of Medicine. 2002;113:308-316.
21. Liu F, Song SC, Mix D, Baudys M, Kim SW. Glucose-induced release of glycosylpoly(ethylene glycol) insulin bound to a soluble conjugate of concavalin A. Bioconjugate Chem. 1997;8:664-672.
22. Wishart DS. Identifying putative drug targets and potential drug leads: starting points for virtual screening and docking. Methods Mol Biol. 2015;1215:425-44.
23. Xie XQ. Exploiting PubChem for Virtual Screening. Expert Opin Drug Discov. 2010;5(12):1205-1220.
24. Irwin JJ, Shoichet BK. ZINC--a free database of commercially available compounds for virtual screening. J Chem Inf Model. 2005;45(1):177-82.
25. Tanabe M, Kanehisa M. Using the KEGG database resource. Curr Protoc Bioinformatics. 2012; Chapter 1: Unit 1.12. doi: 10.1002/0471250953.bi0112s38.
26. Wishart DS, Knox C, Guo AC, Shrivastava S, Hassanali M, Stothard P, Chang Z, Woolsey J. DrugBank: a comprehensive resource for in silico drug discovery and exploration. Nucleic Acids Res. 2006;34(Database issue):D668-72.
27. Regon P, Gogoi D, Rai AK, Bordoloi M, Bezbaruah RL. Design of inhibitors of the HIV-1 integrase core domain using virtual screening. Bioinformation. 2014;10(2):76-80.
28. Papadatos G, van Westen GJ, Croset S, Santos R, Trubian S, Overington JP. A document classifier for medicinal chemistry publications trained on the ChEMBL corpus. J Cheminform. 2014;6(1):40.
29. Hill DP, Adams N, Bada M, Batchelor C, Berardini TZ, Dietze H, Drabkin HJ, Ennis M, Foulger RE, Harris MA, Hastings J, Kale NS, de Matos P, Mungall CJ, Owen G, Roncaglia P, Steinbeck C, Turner S, Lomax J. Dovetailing biology and chemistry: integrating the Gene Ontology with the ChEBI chemical ontology. BMC Genomics. 2013;29;14:513
30. Oprea TI, Davis AM, Teague SJ, Leeson PD. Is There a Difference between Leads and Drugs? A Historical Perspective. J. Chem. Inf. Comput. Sci. 2001;41(5):1308–1315.
31. Zhu X, Ji J, Huang D, Zhu Y, Tang C, Yang X, Qian H, Huang W. Discovery, synthesis and evaluation of novel cholesterol absorption inhibitors. Chem Biol Drug Des. 2012;80(3):426-33.
32. Rybacka A, Rudén C, Andersson PL. On the use of in silico tools for prioritising toxicity testing of the low-volume industrial chemicals in REACH. Basic Clin Pharmacol Toxicol. 2014l;115(1):77-87.
33. Rao SN, Head MS, Kulkarni A, Lalonde JM. Validation Studies of the Site-Directed Docking Program LibDock. J Chem Inf Model. 2007;47(6):2159-2171.
34. Wenjie Liu, Winston Duo Wu, Cordelia Selomulya, Xiao Dong Chen. On designing particulate carriers for encapsulation and controlled release applications. Powder Technology. 2013;236:188–196.
35. Trainor C, Silverman M. Transcapillary exchange of molecular weight markers in the postglomerular circulation: application of a barrier-limited model. Am J Physiol. 1982;242(5):F436-46.

36. Clardy SM, Allis DG, Fairchild TJ, Doyle RP. Vitamin B12 in drug delivery: breaking through the barriers to a B12 bioconjugate pharmaceutical. Expert Opin Drug Deliv. 2011;8(1):127-40.
37. Chalasani KB, Russell-Jones GJ, Jain AK, Diwan PV, Jain SK. Effective oral delivery of insulin in animal models using vitamin B12-coated dextran nanoparticles. J Control Release. 2007a;122(2):141-50.
38. Chalasani KB, Russell-Jones GJ, Yandrapu SK, Diwan PV, Jain SK. A novel vitamin B12-nanosphere conjugate carrier system for peroral delivery of insulin. J Control Release. 2007b;117(3):421-9.
39. Tiwari D, Goldman D, Town C, Sause R, Madan PL. In vitro-in vivo evaluation of a controlled release buccal bioadhesive device for oral drug delivery. Pharm Res. 1999;16(11):1775-80.
40. Petrus AK, Vortherms AR, Fairchild TJ, Doyle RP. Vitamin B12 as a carrier for the oral delivery of insulin. ChemMedChem. 2007;2(12):1717-21.
41. Russell-Jones GJ. Use of vitamin B12 conjugates to deliver protein drugs by the oral route. Crit Rev Ther Drug Carrier Syst. 1998;15(6):557-86.
42. Lildballe DL, Mutti E, Birn H, Nexo E. Maximal load of the vitamin B12 transport system: a study on mice treated for four weeks with high-dose vitamin B12 or cobinamide. PLoS One. 2012;7(10):e46657.
43. Goswami LN, Ma L, Cai Q, Sarma SJ, Jalisatgi SS, Hawthorne MF. cRGD peptide-conjugated icosahedral closo-B12(2-) core carrying multiple Gd3+-DOTA chelates for α(v)β3 integrin-targeted tumor imaging (MRI). Inorg Chem. 2013;52(4):1701-9.
44. Arriaga-Alba M, Ruiz-Pérez NJ, Sánchez-Navarrete J, de Angel BL, Flores-Lozada J, Blasco JL. Antimutagenic evaluation of vitamins B1, B6 and B12 in vitro and in vivo, with the Ames test. Food Chem Toxicol. 2013;53:228-34.
45. Narain R, Gonzales M, Hoffman AS, Stayton PS, Krishnan KM. Synthesis of monodisperse biotinylated p(NIPAAm)-coated iron oxide magnetic nanoparticles and their bioconjugation to streptavidin. Langmuir. 2007;23(11):6299-304.
46. Youn YS, Chae SY, Lee S, Kwon MJ, Shin HJ, Lee KC. Improved peroral delivery of glucagon-like peptide-1 by site-specific biotin modification: design, preparation, and biological evaluation. Eur J Pharm Biopharm. 2008;68(3):667-75.
47. Chae SY, Jin CH, Shin HJ, Youn YS, Lee S, Lee KC. Preparation, characterization, and application of biotinylated and biotin-PEGylated glucagon-like peptide-1 analogues for enhanced oraldelivery. Bioconjug Chem. 2008;19(1):334-41.
48. Jin CH, Chae SY, Son S, Kim TH, Um KA, Youn YS, Lee S, Lee KC. A new orally available glucagon-like peptide-1 receptor agonist, biotinylated exendin-4, displays improved hypoglycemic effects in db/db mice. J Control Release. 2009;133(3):172-7.
49. Newman W, Beall LD, Levine MA, Cone JL, Randhawa ZI, Bertolini DR. Biotinylated parathyroid hormone as a probe for the parathyroid hormone receptor. Structure-function analysis and detection of specific binding to cultured bone cells by flow cytometry. J Biol Chem. 1989;264(28):16359-65.
50. Yang W, Wang M, Ma L, Li H, Huang L. Synthesis and characterization of biotin modified cholesteryl pullulan as a novel anticancer drug carrier. Carbohydr Polym. 2014;99:720-7
51. Song H, Su C, Cui W, Zhu B, Liu L, Chen Z, Zhao L. Folic Acid-chitosan conjugated nanoparticles for improving tumor-targeted drug delivery. Biomed Res Int. 2013;2013:723158.
52. Wang M, Hu H, Sun Y, Qiu L, Zhang J, Guan G, Zhao X, Qiao M, Cheng L, Cheng L, Chen D. A pH-sensitive gene delivery system based on folic acid-PEG-chitosan - PAMAM-plasmid DNA complexes for cancer cell targeting. Biomaterials. 2013;34(38):10120-32.
53. Luo Z, Ding X, Hu Y, Wu S, Xiang Y, Zeng Y, Zhang B, Yan H, Zhang H, Zhu L, Liu J, Li J, Cai K, Zhao Y. Engineering a hollow nanocontainer platform with multifunctional molecular machines for tumor-targeted therapy in vitro and in vivo. ACS Nano. 2013;7(11):10271-84.
54. Ling SS, Yuen KH, Magosso E, Barker SA. Oral bioavailability enhancement of a hydrophilic drug delivered via folic acid-coupled liposomes in rats. J Pharm Pharmacol. 2009;61(4):445-9.
55. Anderson KE, Eliot LA, Stevenson BR, Rogers JA. Formulation and evaluation of a folic acid receptor-targeted oral vancomycin liposomal dosage form. Pharm Res. 2001;18(3):316-22.
56. Yan L, Chen W, Zhu X, Huang L, Wang Z, Zhu G, Roy VA, Yu KN, Chen X. Folic acid conjugated self-assembled layered double hydroxide nanoparticles for high-efficacy-targeted drug delivery. Chem Commun (Camb). 2013;49(93):10938-40.
57. Charernsriwilaiwat N, Rojanarata T, Ngawhirunpat T, Opanasopit P. Electrospun chitosan/polyvinyl alcohol nanofibre mats for wound healing. Int Wound J. 2012;11(2):215-22.
58. Salman HH, Gamazo C, Agüeros M, Irache JM. Bioadhesive capacity and immunoadjuvant properties of thiamine-coated nanoparticles. Vaccine. 2007;25(48):8123-32.
59. DISULFIDE DERIVATIVES OF THIAMINE IN BIOLOGICAL SYSTEMS. Acta Physiologica Scandinavica. 1960;50(s174):10–16.
60. National Academy of Sciences. Institute of Medicine. Food and Nutrition Board., ed. "Chapter 4 - Thiamine". Dietary Reference Intakes for Thiamine, Riboflavin, Niacin, Vitamin B6, Folate, Vitamin B12, Pantothenic Acid, Biotin, and Choline. Washington, D.C.: National Academy Press. pp. 58–86. ISBN 0-309-06411-2, 1998.
61. Sinko PJ, Lee YH, Makhey V, Leesman GD, Sutyak JP, Yu H, Perry B, Smith CL, Hu P, Wagner EJ, Falzone LM, McWhorter LT, Gilligan JP, Stern W. Biopharmaceutical approaches for developing and assessing oral peptide delivery strategies and systems: in vitro permeability and in vivo oral absorption of salmon calcitonin (sCT). Pharm Res. 1999r;16(4):527-33.
62. Matsuda K, Yuasa H, Watanabe J. Fractional absorption of L-carnitine after oral administration in rats: evaluation of absorption site and dose dependency. Biol Pharm Bull. 1998a;21(7):752-5.
63. Matsuda K, Yuasa H, Watanabe J. Physiological mechanism-based analysis of dose-dependent gastrointestinal absorption of L-carnitine in rats. Biopharm Drug Dispos. 1998bt;19(7):465-72.

64. Zhang S, Qin G, Wu Y, Gao Y, Qiu Y, Li F, Xu B. Enhanced bioavailability of L-carnitine after painless intradermal delivery vs. oral administration in rats. Pharm Res. 2011;28(1):117-23.
65. Wagner CC, Rusca A, Kletter K, Tschurlovits M, Pace S, Longo A, Pedrani M, Villa R, Frimonti E, Müller M, Brunner M. Plasma pharmacokinetics and gastrointestinal transit of a new Propionyl-l-Carnitine controlled release formulation. Xenobiotica. 2011;41(11):988-95.
66. Ribas GS, Vargas CR, Wajner M. L-carnitine supplementation as a potential antioxidant therapy for inherited neurometabolic disorders. Gene. 2014;533(2):469-76.
67. Kuskov AN, Voskresenskaya AA, Goryachaya AV, Artyukhov AA, Shtilman MI, Tsatsakis AM. Preparation and characterization of amphiphilic poly-N-vinylpyrrolidone nanoparticles containing indomethacin. J Mater Sci Mater Med. 2010a;21(5):1521-30.
68. Kuskov AN, Voskresenskaya AA, Goryachaya AV, Shtilman MI, Spandidos DA, Rizos AK, Tsatsakis AM. Amphiphilic poly-N-vinylpyrrolidone nanoparticles as carriers for non-steroidal anti-inflammatory drugs: characterization and in vitro controlled release of indomethacin. Int J Mol Med. 2010b;26(1):85-94.
69. Jansen J, Boerakker MJ, Heuts J, Feijen J, Grijpma DW. Rapid photo-crosslinking of fumaric acid monoethyl ester-functionalized poly(trimethylene carbonate) oligomers for drug deliveryapplications. J Control Release. 2010a;147(1):54-61.
70. Kuskov AN, Voskresenskaya AA, Goryachaya AV, Shtilman MI, Spandidos DA, Rizos AK, Tsatsakis AM. Amphiphilic poly-N-vinylpyrrolidone nanoparticles as carriers for non-steroidal anti-inflammatory drugs: characterization and in vitro controlled release of indomethacin. J Mol Med. 2010c;26(1):85-94.
71. Zhu Z, Xie C, Liu Q, Zhen X, Zheng X, Wu W, Li R, Ding Y, Jiang X, Liu B. The effect of hydrophilic chain length and iRGD on drug delivery from poly(ε-caprolactone)-poly(N-vinylpyrrolidone) nanoparticles. Biomaterials. 2011;32(35):9525-35.
72. Jansen J, Bosman MB, Boerakker MJ, Feijen J, Grijpma DW. Photo-crosslinked poly(trimethylene carbonate)-fumarate/n-vinyl pyrrolidone networks for the controlled release of proteins. J Control Release. 2010b;148(1):e79-80.
73. Sajeesh S, Sharma CP. Mucoadhesive hydrogel microparticles based on poly (methacrylic acid-vinyl pyrrolidone)-chitosan for oral drug delivery. Drug Deliv. 2011;18(4):227-35.
74. Baker GA, Campbell ML, Yates MZ, McCleskey TM. Carbon dioxide emulsion assisted loading of polymer microspheres toward sustained release materials. Langmuir. 2005;21(9):3730-2.
75. Pályi I, Vincze B, Lovas S, Mezö I, Pató J, Kálnay A, Turi G, Gaál D, Mihalik R, Péter I, Teplán I, Murphy RF. Gonadotropin-releasing hormone analogue conjugates with strong selective antitumor activity. Proc Natl Acad Sci U S A. 1999;96(5):2361-6.
76. Gomez, A.; Bingham, D.; Juan, L.d.; Tang, K. Production of protein nanoparticles by electrospray drying. Journal of Aerosol Science. 1998;29(5):561-574.
77. Mandracchia D, Denora N, Franco M, Pitarresi G, Giammona G, Trapani G. New Biodegradable Hydrogels Based on Inulin and alpha,beta-Polyaspartylhydrazide Designed for Colonic Drug Delivery: In Vitro Release of Glutathione and Oxytocin. J Biomater Sci Polym Ed. 2011;22(1-3): 313-328.
78. Patel M, Shah T, Amin A. Therapeutic opportunities in colon-specific drug-delivery systems. Crit Rev Ther Drug Carrier Syst. 2007;24(2):147-202.
79. Sinha VR, Kumria R. Polysaccharides in colon-specific drug delivery. Int J Pharm. 2001;224(1-2):19-38.
80. Chourasia MK, Jain SK. Pharmaceutical approaches to colon targeted drug delivery systems. J Pharm Pharm Sci. 2003;6(1):33-66.
81. Ravi V, Siddaramaiah, Pramod Kumar TM. Influence of natural polymer coating on novel colon targeting drug delivery system. J Mater Sci Mater Med. 2008;19(5):2131-6.
82. Proescholdt MG, Hutto B, Brady LS, Herkenham M. Studies of cerebrospinal fluid flow and penetration into brain following lateral ventricle and cisterna magna injections of the tracer [14C]inulin in rat. Neuroscience. 2000;95(2):577-92.
83. Lucchesi KJ, Gosselin RE. Mechanism of L-glucose, raffinose, and inulin transport across intact blood-brain barriers. Am J Physiol. 1990;258(3 Pt 2):H695-705.
84. Kumar S, Tummala H. Development of soluble inulin microparticles as a potent and safe vaccine adjuvant and delivery system. Mol Pharm. 2013;10(5):1845-53.
85. Sokiić ZB, Knezević J, Vrvić MM. Inulin--potential prebiotic. Med Pregl. 2009;62(3-4):153-6.
86. Meyer D, Stasse-Wolthuis M. The bifidogenic effect of inulin and oligofructose and its consequences for gut health. Eur J Clin Nutr. 2009;63(11):1277-89.
87. De Preter V, Ghebretinsae AH, Abrahantes JC, Windey K, Rutgeerts P, Verbeke K. Impact of the synbiotic combination of Lactobacillus casei shirota and oligofructose-enriched inulin on the fecal volatile metabolite profile in healthy subjects. Mol Nutr Food Res. 2011;55(5):714-22.
88. Sherman RL Jr, Ford WT. Semiconductor nanoparticle/polystyrene latex composite materials. Langmuir. 2005;21(11):5218-22.
89. Johnson AM, Holcombe JA. Poly(L-cysteine) as an electrochemically modifiable ligand for trace metal chelation. Anal Chem. 2005;77(1):30-5.
90. Sun J, Chen X, Lu T, Liu S, Tian H, Guo Z, Jing X. Formation of reversible shell cross-linked micelles from the biodegradable amphiphilic diblock copolymer poly(L-cysteine)-block-poly(L-lactide). Langmuir. 2008;24(18):10099-106.
91. Sethuraman VA, Lee MC, Bae YH. A biodegradable pH-sensitive micelle system for targeting acidic solid tumors. Pharm Res. 2008;25(3):657-66.

92. Takatsuka S, Morita T, Horikiri Y, Yamahara H, Saji H. Influence of various combinations of mucolytic agent and non-ionic surfactant on intestinal absorption of poorly absorbed hydrophilic compounds. Int J Pharm. 2008;349(1-2):94-100.
93. Dünnhaupt S, Barthelmes J, Hombach J, Sakloetsakun D, Arkhipova V, Bernkop-Schnürch A. Distribution of thiolated mucoadhesive nanoparticles on intestinal mucosa. Int J Pharm. 201;408(1-2):191-9.
94. Licciardi M, Pitarresi G, Cavallaro G, Giammona G. Nanoaggregates based on new poly-hydroxyethyl-aspartamide copolymers for oral insulin absorption. Mol Pharm. 2013;10(5):1644-54.
95. Iqbal J, Shahnaz G, Dünnhaupt S, Müller C, Hintzen F, Bernkop-Schnürch A. Preactivated thiomers as mucoadhesive polymers for drug delivery. Biomaterials. 2012;33(5):1528-35.
96. Pashkunova-Martic I, Kremser C, Galanski M, Arion V, Debbage P, Jaschke W, Keppler B. Lectin-Gd-loaded chitosan hydrogel nanoparticles: a new biospecific contrast agent for MRI. Mol Imaging Biol. 2011;13(1):16-24.
97. Arteche Pujana M, Pérez-Álvarez L, Cesteros Iturbe LC, Katime I. Biodegradable chitosan nanogels crosslinked with genipin. Carbohydr Polym. 2013;94(2):836-42.
98. Chang YC, Chen DH. Preparation and adsorption properties of monodisperse chitosan-bound Fe3O4 magnetic nanoparticles for removal of Cu(II) ions. Colloid Interface Sci. 2005 Mar 15;283(2):446-51.
99. Shelma R, Sharma CP. Submicroparticles composed of amphiphilic chitosan derivative for oral insulin and curcumin release applications. Colloids Surf B Biointerfaces. 2011;88(2):722-8.
100. Elsayed A, Al-Remawi M, Qinna N, Farouk A, Al-Sou'od KA, Badwan AA. Chitosan-sodium lauryl sulfate nanoparticles as a carrier system for the in vivo delivery of oral insulin. AAPS PharmSciTech. 2011;12(3):958-64.
101. Makhlof A, Tozuka Y, Takeuchi H. Design and evaluation of novel pH-sensitive chitosan nanoparticles for oral insulin delivery. Eur J Pharm Sci. 2011;42(5):445-51.
102. Nagpal K, Singh SK, Mishra DN. Chitosan nanoparticles: a promising system in novel drug delivery. Chem Pharm Bull (Tokyo). 2010;58(11):1423-30.
103. Cadete A, Figueiredo L, Lopes R, Calado CC, Almeida AJ, Gonçalves LM. Development and characterization of a new plasmid delivery system based on chitosan-sodium deoxycholate nanoparticles. Eur J Pharm Sci. 2012;45(4):451-8.
104. Lee H, Jeong C, Ghafoor K, Cho S, Park J. Oral delivery of insulin using chitosan capsules cross-linked with phytic acid. Biomed Mater Eng. 2011;21(1):25-36.
105. Fang A, Cathala B. Smart swelling biopolymer microparticles by a microfluidic approach: synthesis, in situ encapsulation and controlled release. Colloids Surf B Biointerfaces. 2011;82(1):81-6.
106. Sriamornsak P. Application of pectin in oral drug delivery. Expert Opin Drug Deliv. 2011;8(8):1009-23.
107. Wong TW, Colombo G, Sonvico F. Pectin matrix as oral drug delivery vehicle for colon cancer treatment. AAPS PharmSciTech. 2011;12(1):201-14.
108. Patel M, Shah T, Amin A. Therapeutic opportunities in colon-specific drug-delivery systems. Crit Rev Ther Drug Carrier Syst. 2007;24(2):147-202.
109. Fan L, Sun Y, Xie W, Zheng H, Liu S. Oxidized Pectin Cross-Linked Carboxymethyl Chitosan: A New Class of Hydrogels. J Biomater Sci Polym Ed. 2011; 23(16):2119–2132.
110. Si L, Zhao Y, Huang J, Li S, Zhai X, Li G. Calcium pectinate gel bead intended for oral protein delivery: preparation improvement and formulation development. Chem Pharm Bull (Tokyo). 2009;57(7):663-7.
111. Das S, Chaudhury A, Ng KY. Polyethyleneimine-modified pectin beads for colon-specific drug delivery: in vitro and in vivo implications. J Microencapsul. 2011;28(4):268-79.
112. Yu SS, Scherer RL, Ortega RA, Bell CS, O'Neil CP, Hubbell JA, Giorgio TD. Enzymatic- and temperature-sensitive controlled release of ultrasmall superparamagnetic iron oxides (USPIOs). J Nanobiotechnology. 2011;9:7.
113. Buckwalter DJ, Hudson AG, Moore RB, Long TE. Synthesis and characterization of poly(propylene glycol) polytrioxamide and poly(urea oxamide) segmented copolymers. Polym. Int. 2013;63(7);1184–1191.
114. Andreopoulos AG. Plasticization of biodegradable polymers for use in controlled release. Clin Mater. 1994;15(2):89-92.
115. Urbanetz NA. Stabilization of solid dispersions of nimodipine and polyethylene glycol 2000. Eur J Pharm Sci. 2006;28(1-2):67-76.
116. Yeh MK, Davis SS, Coombes AG. Improving protein delivery from microparticles using blends of poly(DL lactide co-glycolide) and poly(ethylene oxide)-poly(propylene oxide) copolymers. Pharm Res. 1996;13(11):1693-8.
117. Croy SR, Kwon GS. Polymeric micelles for drug delivery. Curr Pharm Des. 2006;12(36):4669-84.
118. Final Report on the Safety Assessment of Propylene Glycol and Polypropylene Glycols. International Journal of Toxicology. 1994;13(6):437-491.
119. Final Report of the Safety Assessment of Propylene Glycol and Polypropylene Glycols, Cosmetic Ingredient Review. J. Am. Coll. Toxicol. 1994;13(6):473-91.
120. Hoenig V, Werner F. Is propylene glycol an inert substance? Toxicol Lett. 1980;5(6):389-92.
121. Funk JO, Maibach HI. Propylene Glycol Dermatitis: Re -evaluation of an Old Problem. Contact Dermatitis. 1994;31(2):36-41
122. Li G, Li P, Qiu H, Li D, Su M, Xu K. Synthesis, characterizations and biocompatibility of alternating block polyurethanes based on P3/4HB and PPG-PEG-PPG. J Biomed Mater Res A. 2011;98(1):88-99.
123. Seidenfeld MA, Hanzlik PJ. The general properties, actions, and toxicity of propylene glycol. J Pharmacol. 1932;44:109.
124. Martin G, Finberg L. Propylene glycol: a potentially toxic vehicle in liquid dosage form. J Pediatr. 1970;77(5):877-8.

125. Glasgow AM, Boeckx RL, Miller MK, MacDonald MG, August GP, Goodman SI. Hyperosmolality in small infants due to propylene glycol. Peaditrics. 1983;72(3):353-5.
126. Runge MB, Wang H, Spinner RJ, Windebank AJ, Yaszemski MJ. Reformulating polycaprolactone fumarate to eliminate toxic diethylene glycol: effects of polymeric branching and autoclave sterilization on material properties. Acta Biomater. 2012;8(1):133-43.
127. IPCS Chemical Safety Information from Intergovernmental Organizations from www.inchem.org: Polypropylene glycol: Hazardous characteristics, 1994; Available from: www.inchem.org
128. Chen CZ, Beck-Tan NC, Dhurjati P, van Dyk TK, LaRossa RA, Cooper SL. Quaternary ammonium functionalized poly(propylene imine) dendrimers as effective antimicrobials: structure-activity studies. Biomacromolecules. 2000;1(3):473-80.
129. Floriano PN, Noble CO 4th, Schoonmaker JM, Poliakoff ED, McCarley RL. Cu(0) nanoclusters derived from poly(propylene imine) dendrimer complexes of Cu(II). J Am Chem Soc. 2001;123(43):10545-53.
130. Paleos CM, Tsiourvas D, Sideratou Z, Tziveleka L. Acid- and salt-triggered multifunctional poly(propylene imine) dendrimer as a prospective drug delivery system. Biomacromolecules. 2004;5(2):524-9.
131. Chooi KW, Gray AI, Tetley L, Fan Y, Uchegbu IF. The molecular shape of poly(propylenimine) dendrimer amphiphiles has a profound effect on their self assembly. Langmuir. 2010;26(4):2301-16.
132. Sideratou Z, Kontoyianni C, Drossopoulou GI, Paleos CM. Synthesis of a folate functionalized PEGylated poly(propylene imine) dendrimer as prospective targeted drug delivery system. Bioorg Med Chem Lett. 2010;20(22):6513-7.
133. Xu Q, Xia Y, Wang CH, Pack DW. Monodisperse double-walled microspheres loaded with chitosan-p53 nanoparticles and doxorubicin for combined gene therapy and chemotherapy. J Control Release. 2012;163(2):130-5.
134. Arnold MM, Gorman EM, Schieber LJ, Munson EJ, Berkland C. NanoCipro encapsulation in monodisperse large porous PLGA microparticles. J Control Release. 2007;121(1-2):100-9.
135. Fattahi P, Borhan A, Abidian MR. Microencapsulation of chemotherapeutics into monodisperse and tunable biodegradable polymers via electrified liquid jets: control of size, shape, and drug release. Adv Mater. 2013;25(33):4555-60.
136. Lee JY, Lee SH, Oh MH, Kim JS, Park TG, Nam YS. Prolonged gene silencing by siRNA/chitosan-g-deoxycholic acid polyplexes loaded within biodegradable polymer nanoparticles. J Control Release. 2012;162(2):407-13.
137. Jungraithmayr W, Laube I, Hild N, Stark WJ, Mihic-Probst D, Weder W, Buschmann J. Bioactive nanocomposite for chest-wall replacement: Cellular response in a murine model. J Biomater Appl. 2013; 29(1):36-45.
138. Wischke C, Zhang Y, Mittal S, Schwendeman SP. Development of PLGA-based injectable delivery systems for hydrophobic fenretinide. Pharm Res. 2010;27(10):2063-74.
139. Pan Y, Xu H, Zhao HY, Wei G, Zheng JM. Study on preparation and oral efficacy of insulin-loaded poly(lactic-co-glycolic acid) nanoparticles. Yao Xue Xue Bao. 2002;37(5):374-7.
140. Kollipara S, Bende G, Movva S, Saha R. Application of rotatable central composite design in the preparation and optimization of poly(lactic-co-glycolic acid) nanoparticles for controlled delivery of paclitaxel. Drug Dev Ind Pharm. 2010;36(11):1377-87.
141. Santhanalakshmi J, Shanthalakshmi G, Aswal VK, Goyal PS. Studies on the microstructure of bile salt aggregates in aqueous n-alkanol solutions by Small Angle Neutron Scattering. Indian J Biochem Biophys. 2002;39(3):170-8.
142. Wang F, Chen Y, Zhang D, Zhang Q, Zheng D, Hao L, Liu Y, Duan C, Jia L, Liu G. Folate-mediated targeted and intracellular delivery of paclitaxel using a novel deoxycholic acid-O-carboxymethylated chitosan-folic acid micelles. Int J Nanomedicine. 2012;7:325-37.
143. Bhattacharjee J, Verma G, Aswal VK, Date AA, Nagarsenker MS, Hassan PA. Tween 80-sodium deoxycholate mixed micelles: structural characterization and application in doxorubicin delivery. J Phys Chem B. 2010;114(49):16414-21.
144. Kim SK, Lee S, Jin S, Moon HT, Jeon OC, Lee DY, Byun Y. Diabetes correction in pancreatectomized canines by orally absorbable insulin-deoxycholate complex. Mol Pharm. 2010;7(3):708-17.
145. Lee S, Lee J, Lee DY, Kim SK, Lee Y, Byun Y. A new drug carrier, Nalpha-deoxycholyl-L: -lysyl-methylester, for enhancing insulin absorption in the intestine. Diabetologia. 2005;48(3):405-11.
146. Qiao D, Gaitonde SV, Qi W, Martinez JD. Deoxycholic acid suppresses p53 by stimulating proteasome-mediated p53 protein degradation. Carcinogenesis. 2001;22(6):957-64.
147. Menting JG, Whittaker J, Margetts MB, Whittaker LJ, Kong GK, Smith BJ, Watson CJ, Záková L, Kletvíková E, Jiráček J, Chan SJ, Steiner DF, Dodson GG, Brzozowski AM, Weiss MA, Ward CW, Lawrence MC. How insulin engages its primary binding site on the insulin receptor. Nature. 2013;493(7431):241-5.
148. Knudsen L, De Meyts P, Kiselyov VV. Insight into the molecular basis for the kinetic differences between the two insulin receptor isoforms. Biochem J. 2011;440(3):397-403.
149. Smith BJ, Huang K, Kong G, Chan SJ, Nakagawa S, Menting JG, Hu SQ, Whittaker J, Steiner DF, Katsoyannis PG, Ward CW, Weiss MA, Lawrence MC. Structural resolution of a tandem hormone-binding element in the insulin receptor and its implications for design of peptide agonists. Proc Natl Acad Sci U S A. 2010;107(15):6771-6.
150. Whittaker J, Whittaker LJ, Roberts CT Jr, Phillips NB, Ismail-Beigi F, Lawrence MC, Weiss MA. α-Helical element at the hormone-binding surface of the insulin receptor functions as a signaling element to activate its tyrosine kinase. Proc Natl Acad Sci U S A. 2012;109(28):11166-71.



Flowchart 1: Screening process of Drug delivering molecules

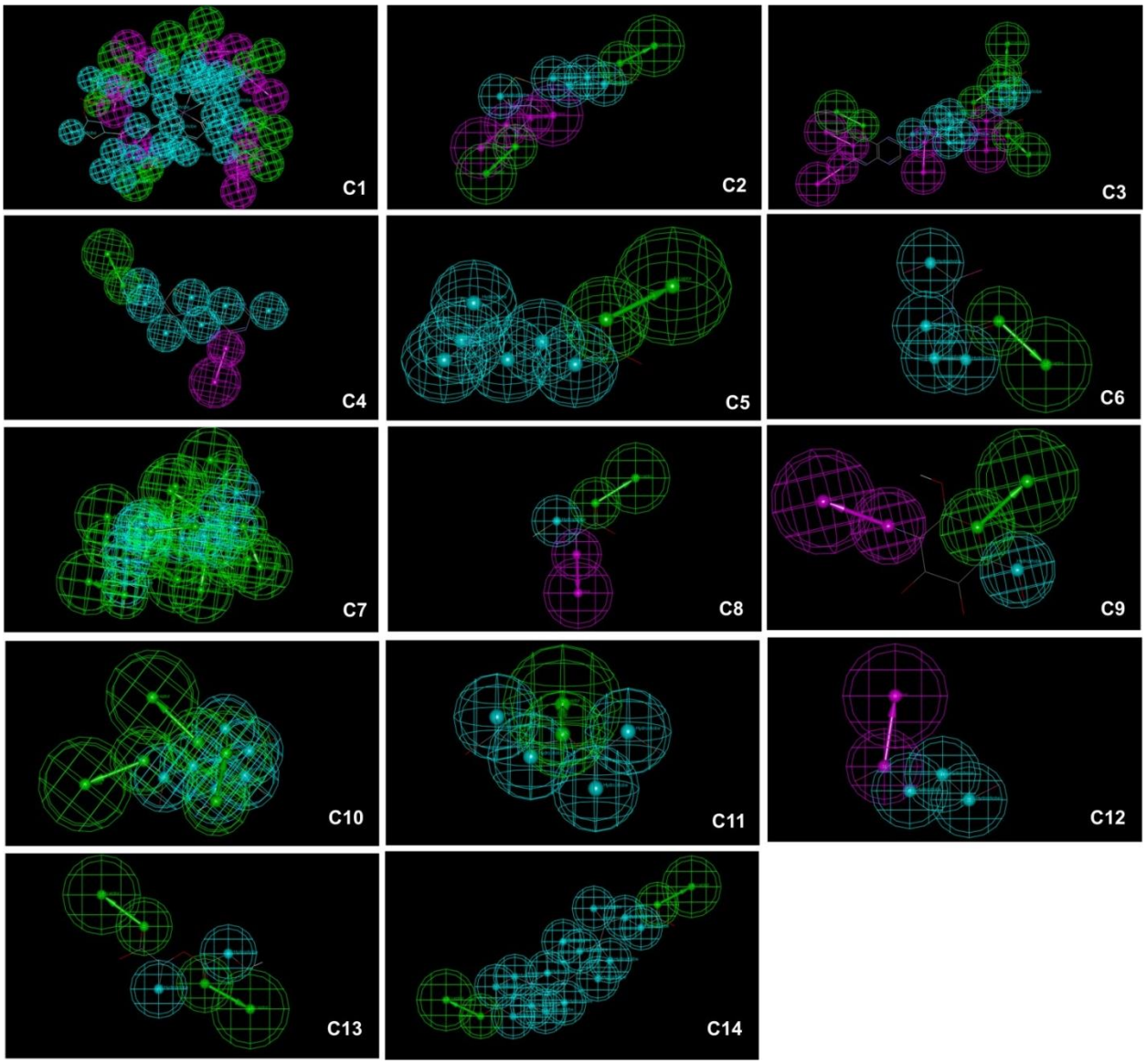


Figure 1

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

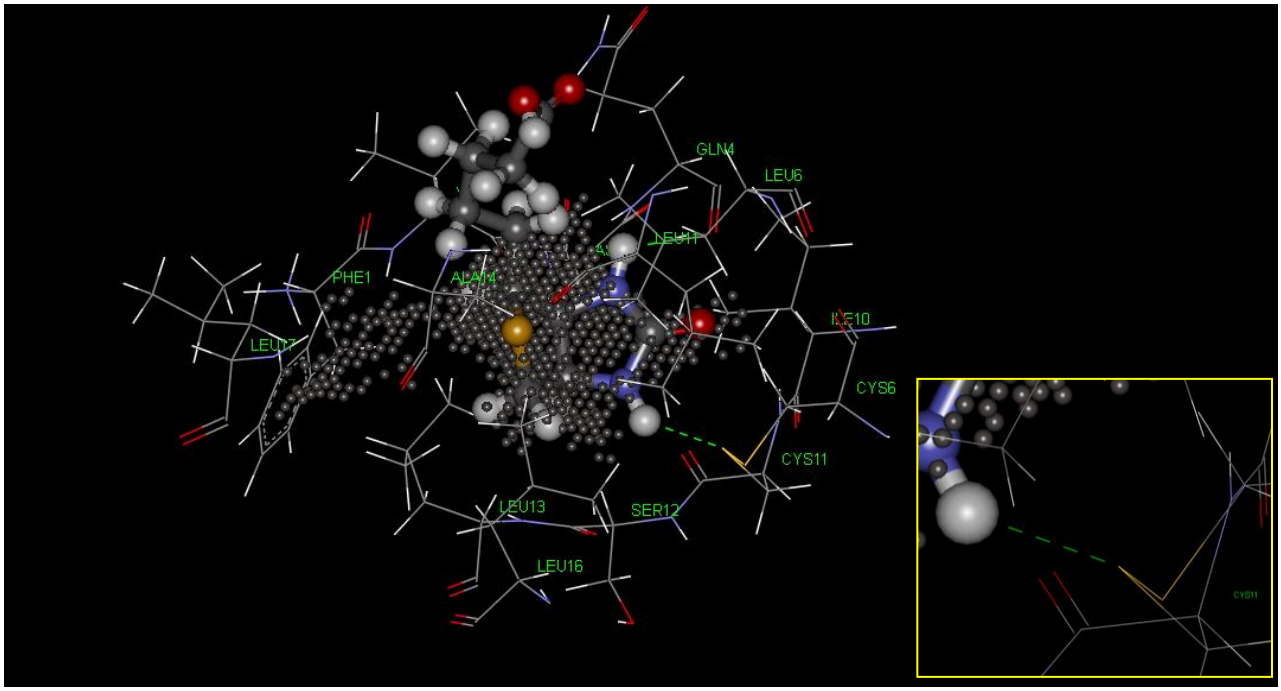


Figure 2a
Bioconjugate from Human Insulin Monomer (PDB ID: 1HLS) and Vitamin H is illustrated by Discovery Studio software. Vitamin H conjugated at CYS A11, GLN B4 and HIS B10 aminoacids of Insulin Monomer. The inner figure illustrates the CYS 11 amino acid of Insulin-A chain conjugates with Vitamin H.

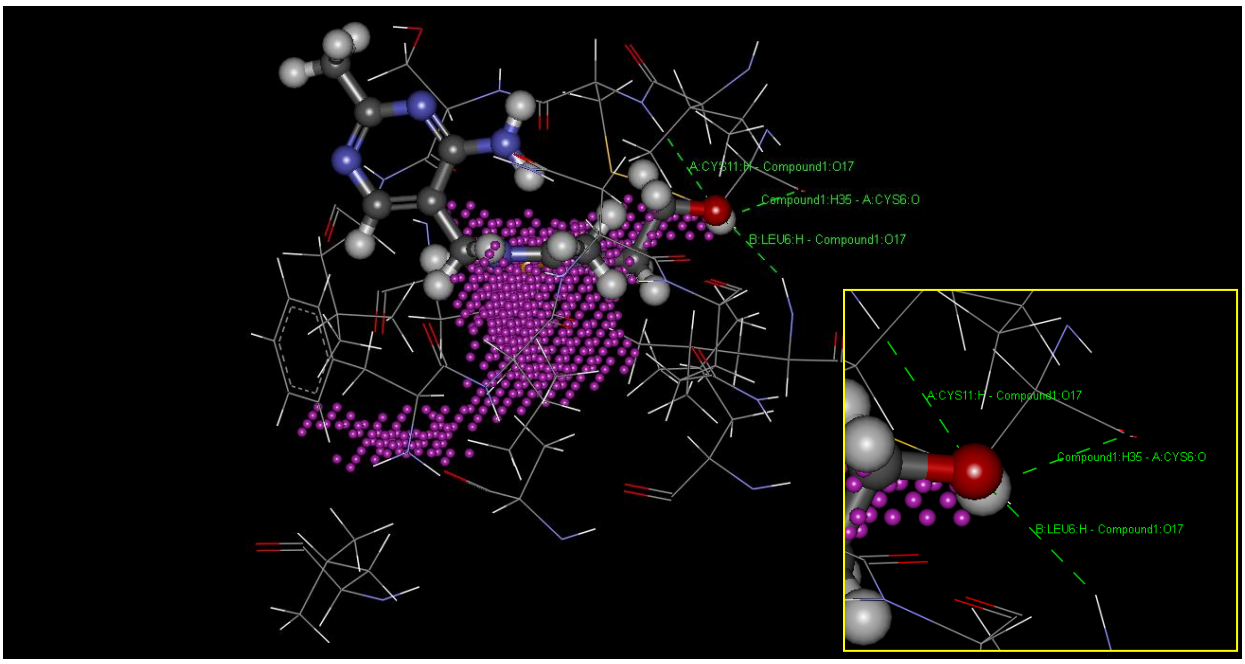


Figure 2b
Bioconjugate from Human Insulin Monomer (PDB ID: 1HLS) and Vitamin B1 is illustrated by Discovery Studio software. Vitamin B1 conjugated at CYS A6, CYS A11 and LEU B6 aminoacids of Insulin Monomer. The inner figure illustrates the CYS 6, CYS 11 amino acids of Insulin-A chain and LEU 6 amino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

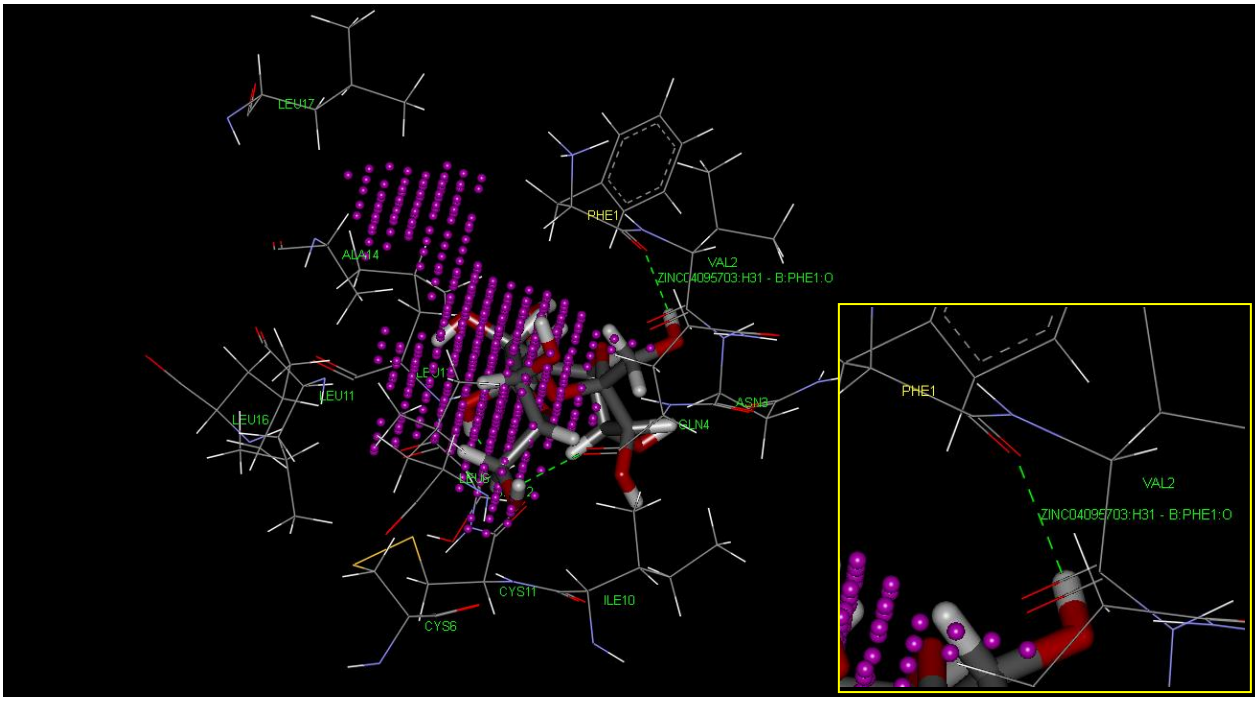


Figure 2c

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.

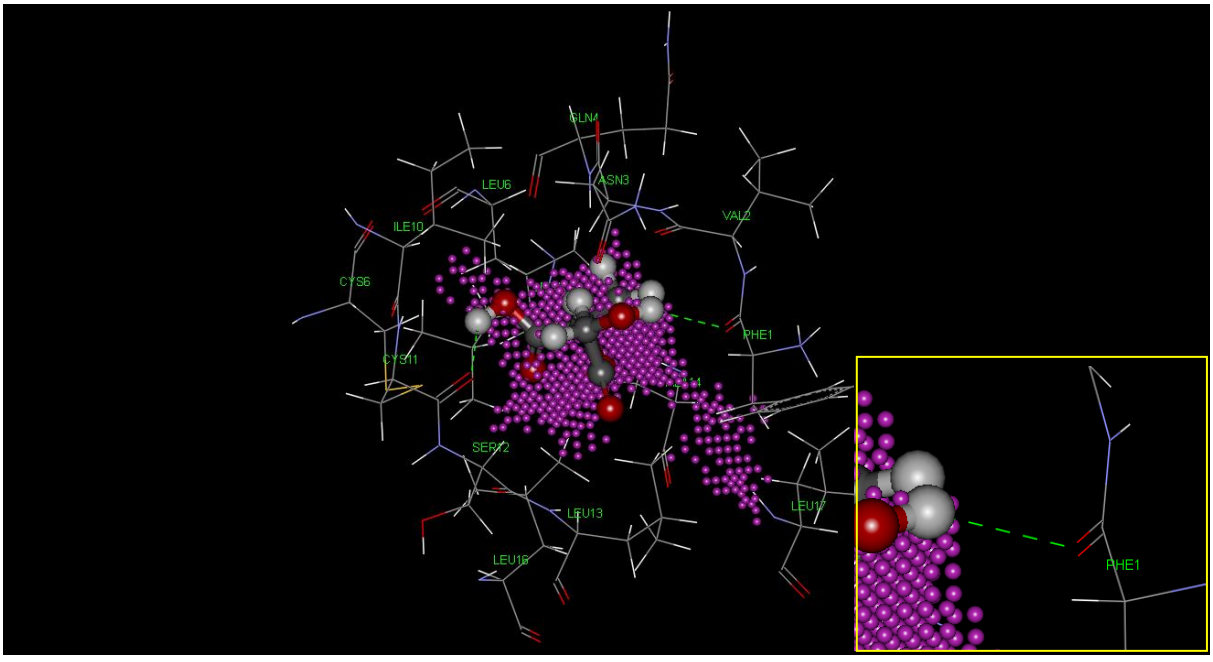


Figure 2d

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).

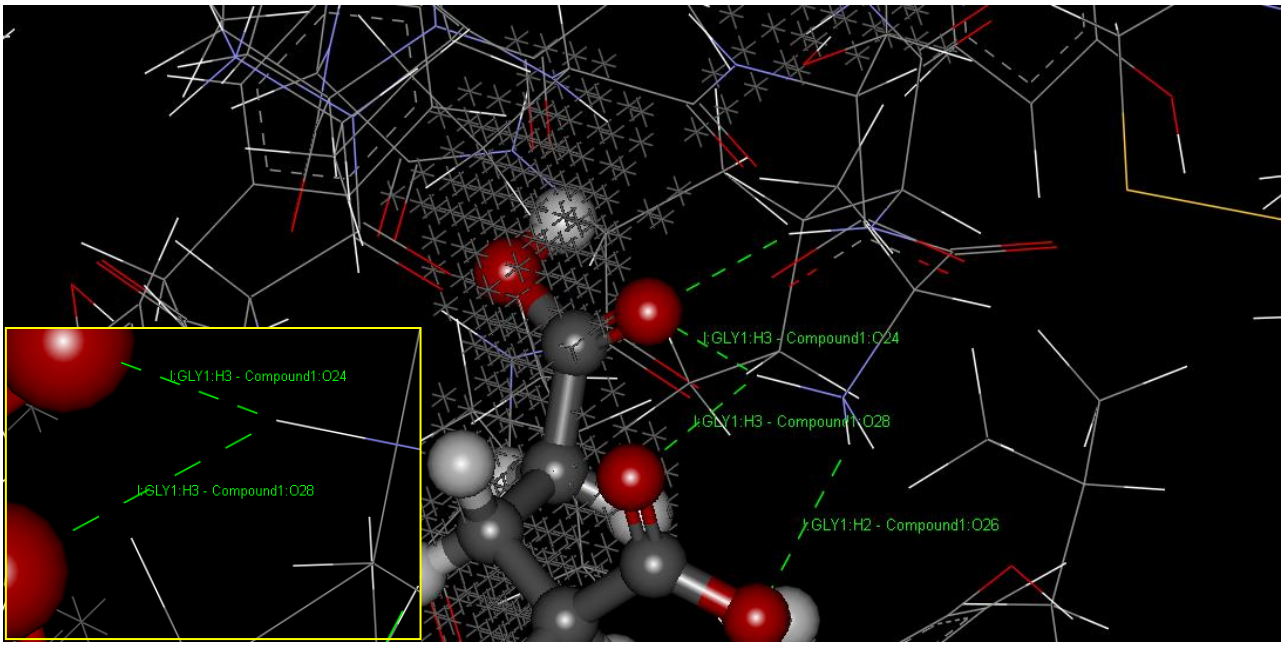


Figure 3a
Bioconjugate from Human insulin hexamer (PDB ID: 1AIO) and Vitamin M is illustrated by Discovery Studio software. Vitamin M conjugated at GLY A1 and THR B27 aminoacids of Insulin hexamer. The inner figure illustrates the GLY 1 amino acid of Insulin-A chain conjugates with Vitamin M.

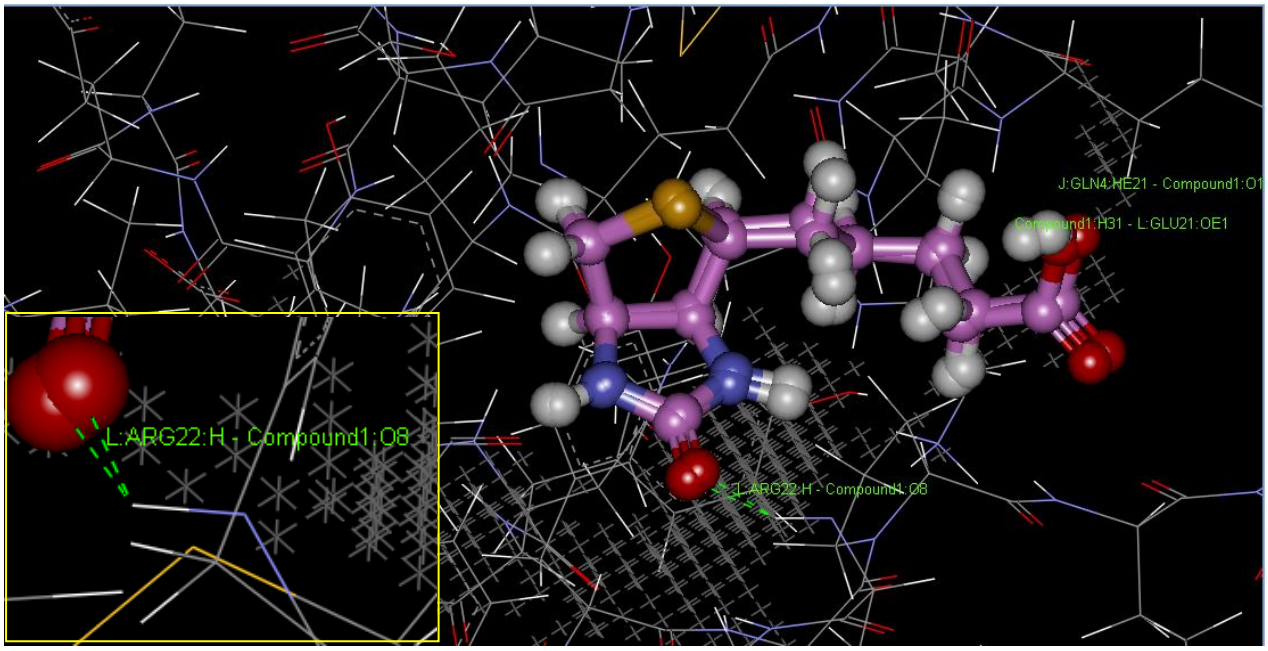


Figure 3b
Bioconjugate from Human insulin hexamer (PDB ID: 1AIO) and Vitamin H is illustrated by Discovery Studio software. Vitamin H conjugated at GLN B4 and ARG B22 aminoacids of Insulin hexamer. The inner figure illustrates the ARG 22 amino acid of Insulin-B chain conjugates with Vitamin H.

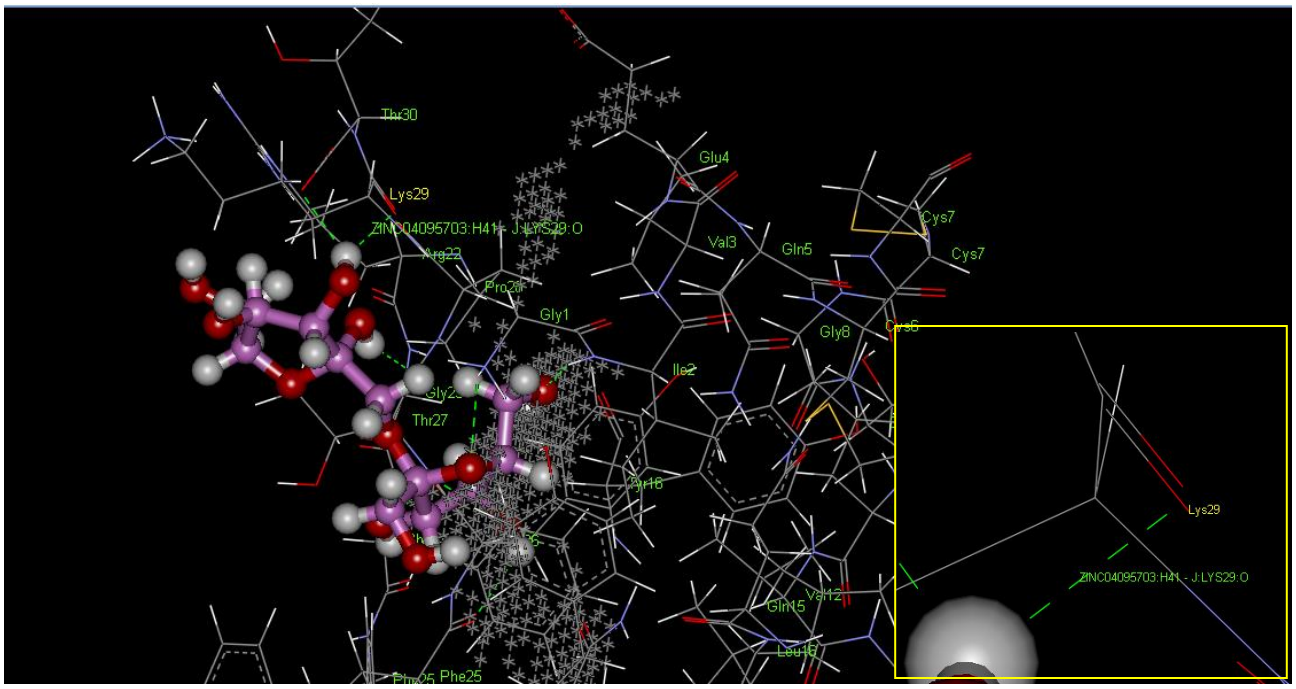


Figure 3c

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.

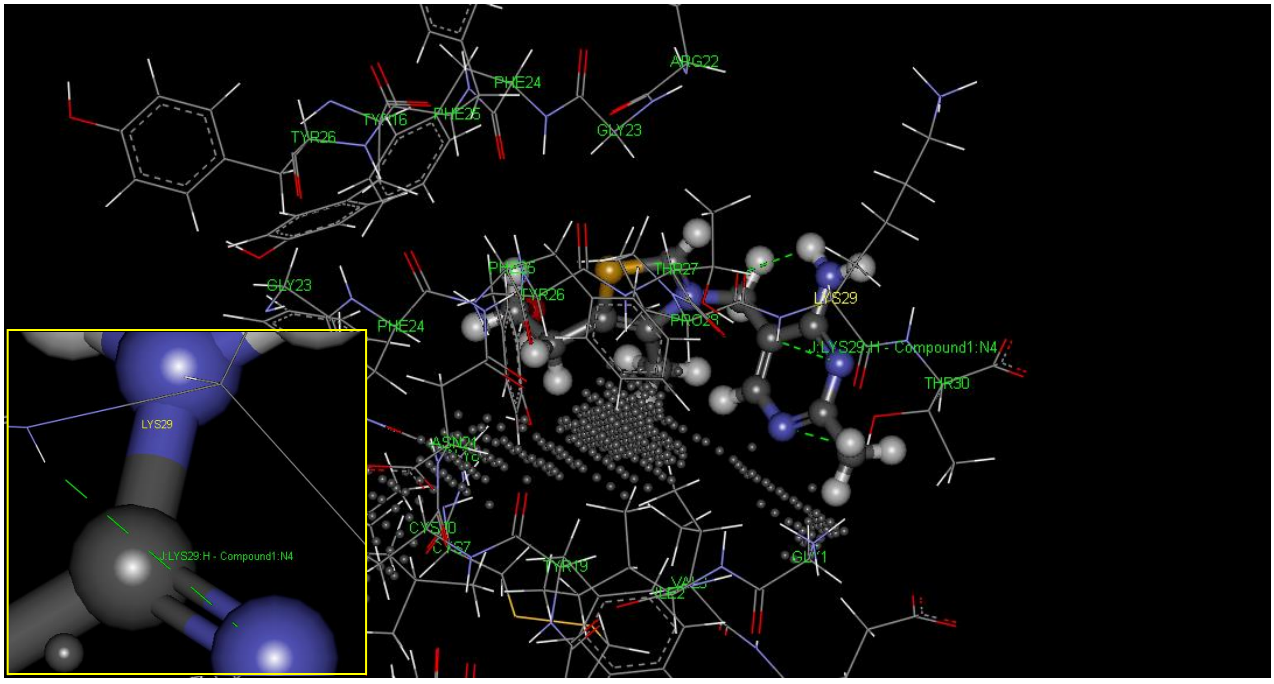


Figure 3d

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.

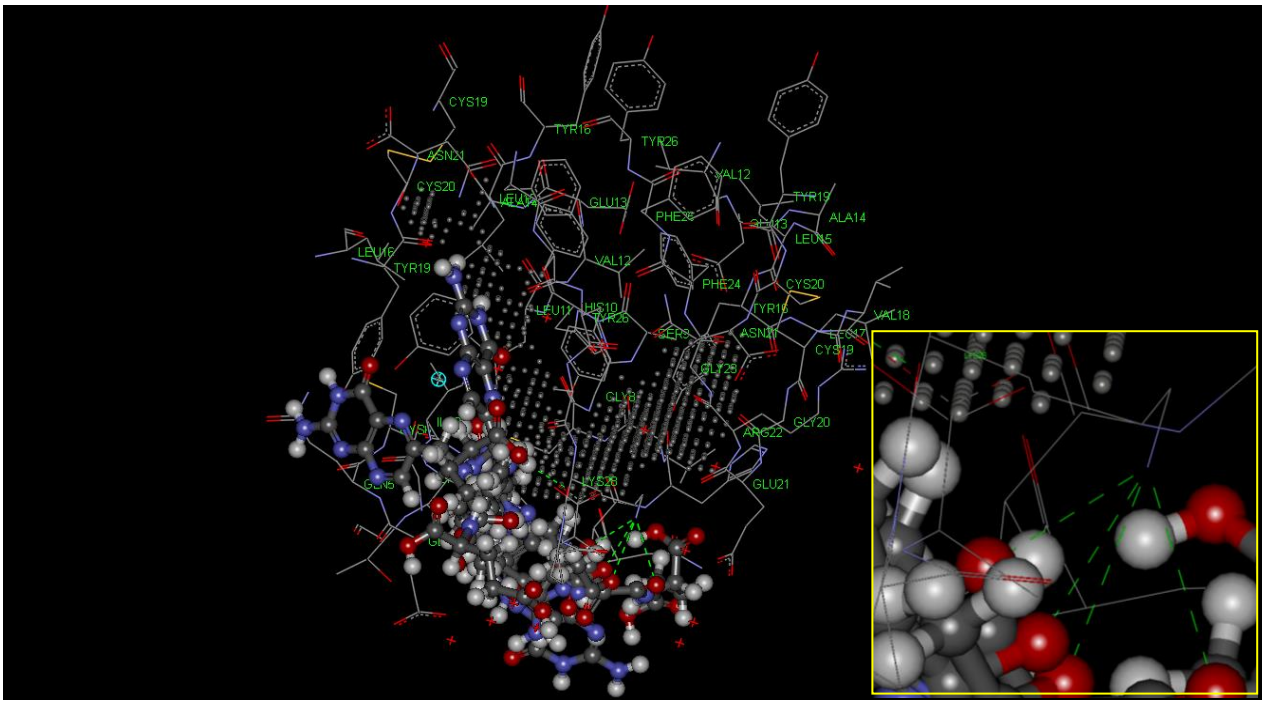


Figure 4a

Bioconjugate from Insulin Lispro (PDB ID: 1 LPH) 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.

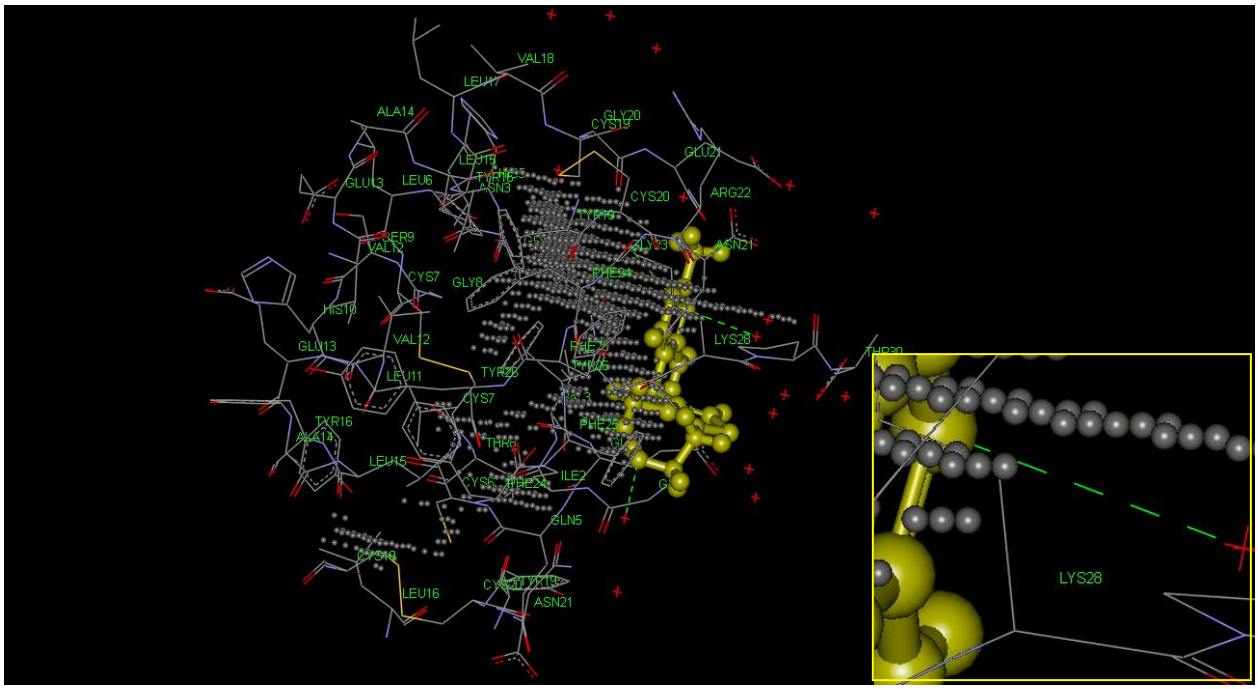


Figure 4b

Bioconjugate from Insulin Lispro (PDB ID: 1 LPH) 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.

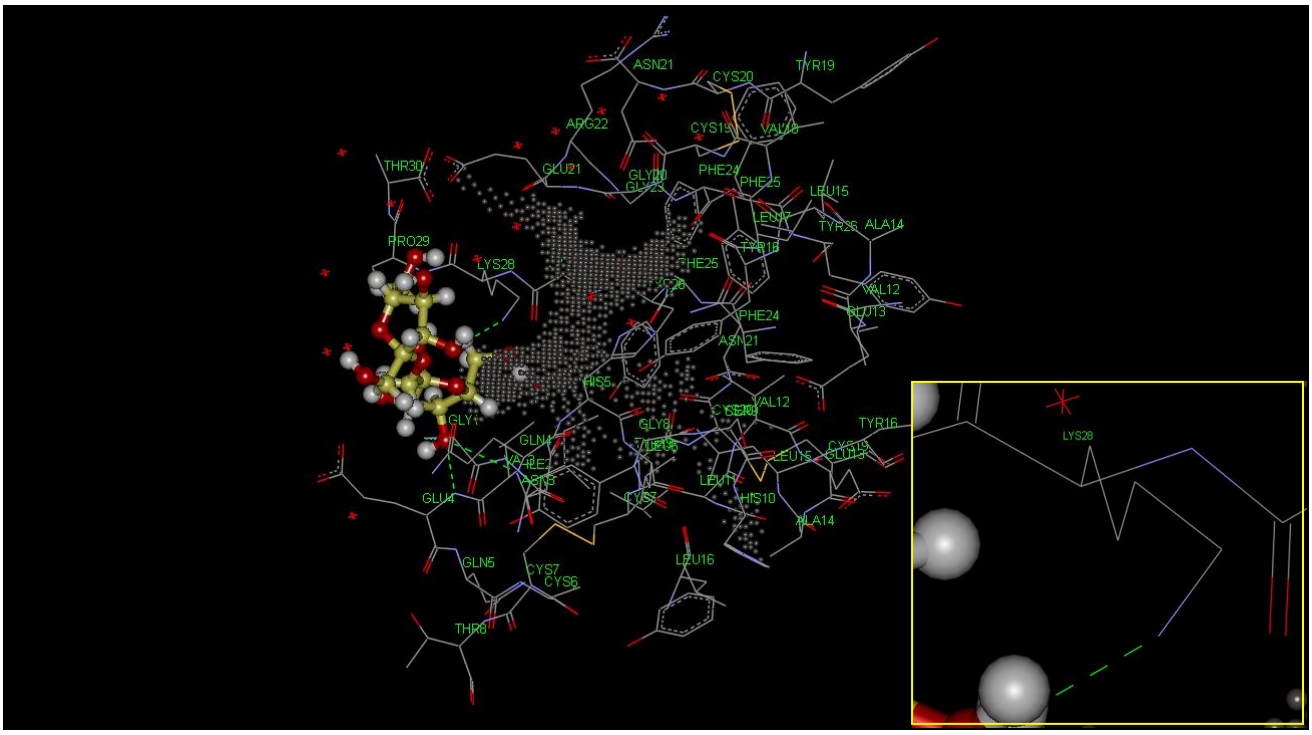


Figure 4c

Bioconjugate from Insulin Lispro (PDB ID: 1 LPH) 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.

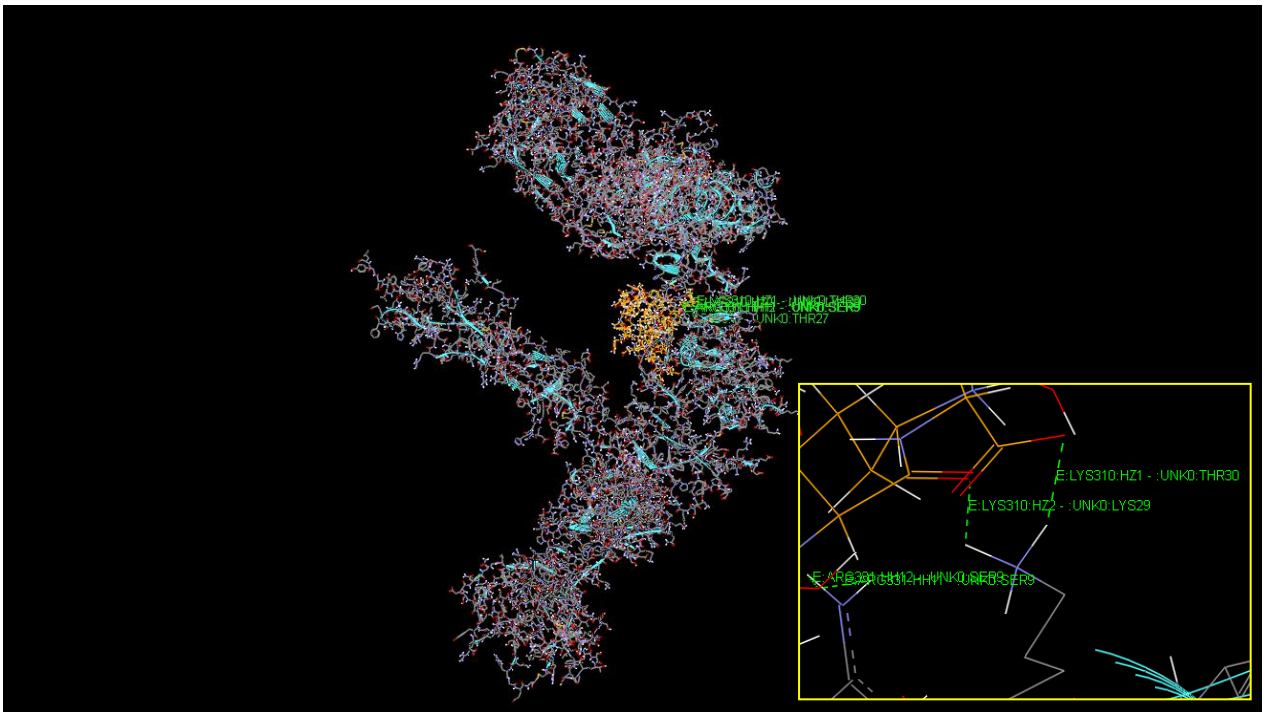


Figure 5a

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 therapeutic effect.

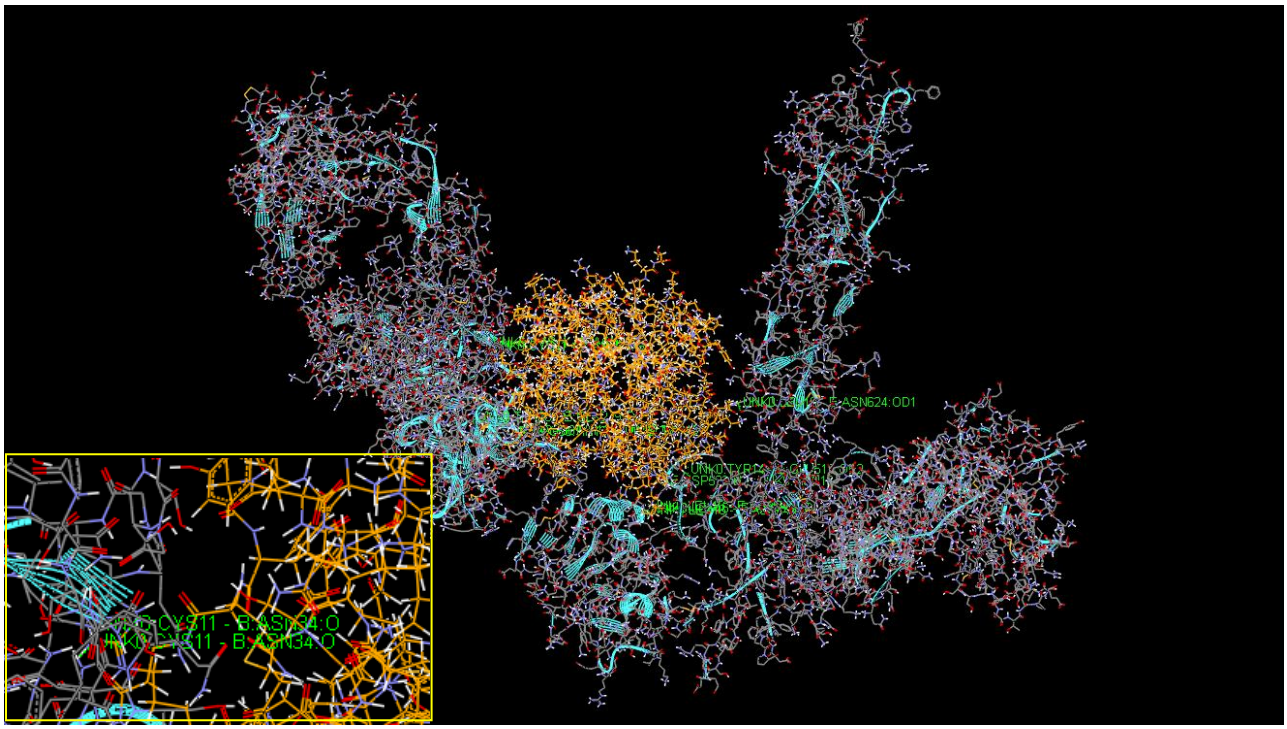


Figure 5b

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). The inner figure shows the interaction in ASN34 aminoacid of IR which is responsible for initiation of therapeutic effect.

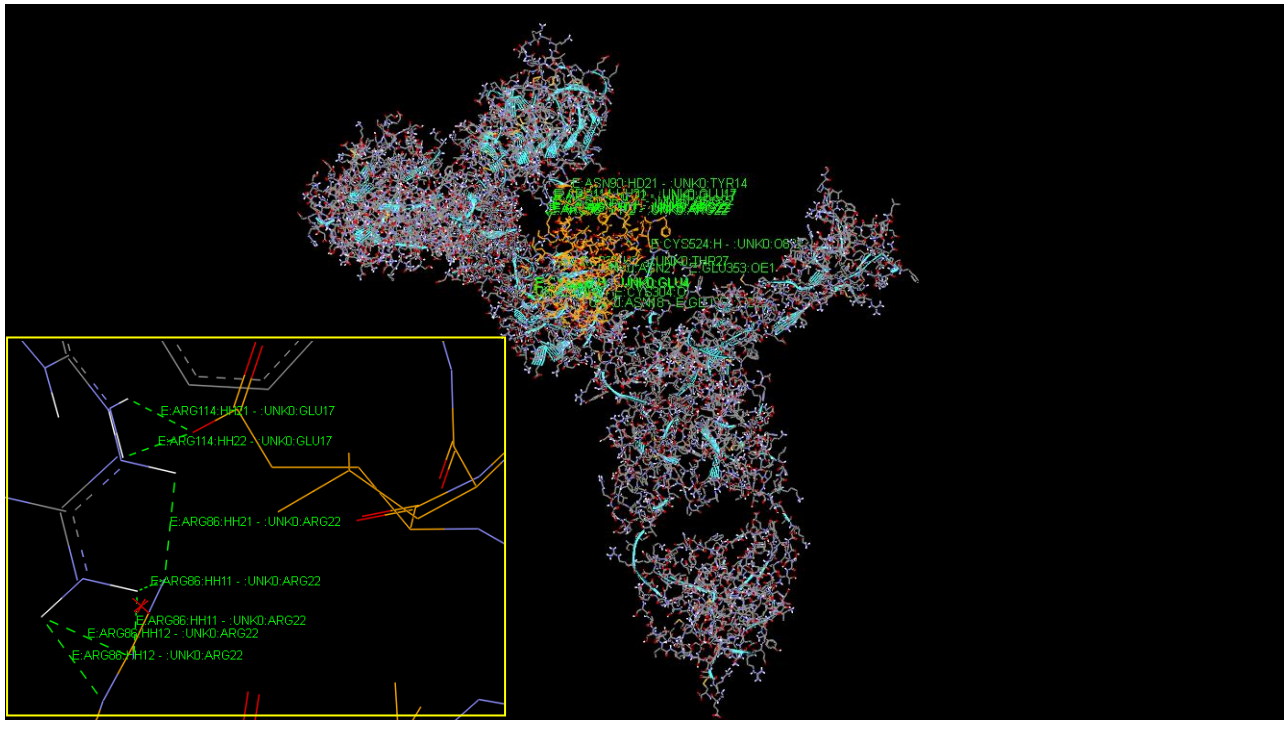


Figure 5c

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 1 Physco chemical details of DDMs by using experimental text mining & filtration criteria

S No	Carrier	Database & Identification Number	Filtration Criteria by experimental text mining						
			Is it Mono-disperse while participate in Drug delivering Conjugate?	Is it Short-chain (low-molecular weight (<2000Da))? (in Daltons)	Is it Biocompatible & Biodegradable?	Is it lipophilic? (Bioavailability ; Efficient intestinal transport)	Is it physically stability against gastric acids and proteolytic enzymes?	Is it inert while participate in Drug delivering Conjugate (no biological activity)?	Is it non-toxic?
C1	Vitamin B12 (cobalamin)	CHEBI ID:30411	Mono-disperse[34]	1,357 [35]	Biocompatible & Biodegradable[36-38]	Oral bioavailability is high[36-39]	Stable[37,38,40,41]	Inert[42]	Non–toxic & Anti mutagenic[43,44]
C2	Vitamin H (Biotin)	CHEBI ID:15956	Mono-disperse[45]	244	Biocompatible & Biodegradable[46-48]	Efficient intestinal transport[46-48]	Stable[46-48]	Inert[49]	Non –toxic[50]
C3	Folic acid (Vitamin M / Vitamin B9)	CHEBI ID:27470	Mono-disperse[51]	441.40	Biocompatible & Biodegradable[52,53]	Oral bioavailability is high[54,55]	Stable[54,55]	No literature	Non –toxic[54-56]
C4	Vitamin B1 (Thiamin)	CHEMBL ID: 1547	No literature	265.112	Biocompatible & Biodegradable[57]	Oral bioavailability is high[58]	Stable[58]	Inert[59]	Non–toxic[60]
C5	L-Carnitine (Vitamin BT)	CHEMBL ID: 1149	No literature	162.113	Biocompatible & Biodegradable[61-64]	Efficient intestinal transport[61-65]	Stable[61-65]	No literature	Non –toxic[66]
C6	Poly-N-vinylpyrrolidone	Pubchem Compound ID: 6917	Mono-disperse[67-68]	11.141	Biocompatible & Biodegradable[69]	Oral bioavailability is low[70,71]	Stable[72-73]	Inert[74]	Moderately toxic[75]
C7	Inulin	Zinc ID: 12358861	Mono-disperse[76]	342.297	Biocompatible & Biodegradable[77]	Oral bioavailability is high[78]	Stable[79-81]	Inert[82,83]	Non –toxic[84-87]
C8	Poly Cysteine	DrugBank ID: DB00151	Mono-disperse[88]	121.158 [89]	Biocompatible & Biodegradable[90,91]	Hydrophilic; poor intestinal absorption [92,93]	Stable in the presence of the peptidase α-chymotrypsin, increase insulin stability against	No literature	Non –toxic[95]

							enzymatic degradation [94]		
C9	Chitosan	ChemSpider ID: 2342878	Mono-disperse[96-98]	501.482	Biocompatible & Biodegradable[97,99-103]	intestinal permeability is Moderate[99-104]	Stable[99-102]	Inert[99,102]	Non – toxic[99,102]
C10	Pectin	KEGG ID: C00714	Mono-disperse[105]	194.139	Biocompatible & Biodegradable (Blood Compatible)[106-109]	intestinal permeability is Moderate[108]	Stable (Colon specific delivery)[110,111]	No literature	Non – toxic[109]
C11	Poly(propylene glycol)	CHEBI ID:53262	Mono-disperse[112,113]	76.094	Biocompatible & Biodegradable[114]	Oral bioavailability is high[115-116]	Stable (protection of encapsulated substances from degradation)[117]	Inert[118-122]	Moderately Toxic[123-125] Less toxic than PEG[126,127]
C12	Poly(propylene imine)	CHEBI ID: 53266	Mono-disperse[128-129]	Molecular weight varies based on length of chain	Biocompatible & Biodegradable[130]	Lipophilic[131]	Stable[130]	No literature	Moderately Toxic[132]
C13	Poly (lactic-co-glycolic acid)	CHEBI:53493	Mono-disperse[133,134]	Molecular weight varies based on length of chain	Biocompatible & Biodegradable[135-137]	Lipophilic[138,139]	Stable[136-139]	Inert[137]	Non – toxic[140]
C14	Deoxycholic acid	DrugBank ID: DB03619	Mono-disperse[141]	392.57	Biocompatible & Biodegradable[142,143]	Lipophilic[144,145]	Stable[144,145]	Not inert[144,145]	Moderately Toxic[146]

Table 2
Pharmacophoric features of carriers

Carrier No	**Carrier**	Number of Acceptor in molecule	Number of Donors in molecule	Number of hydrophobic Region	Number of Pharmacophore principle
C1	Vitamin B12 (cobalamin)	8	7	29	Nil
C2	Vitamin H (Biotin)	2	2	5	6
C3	*Folic acid* (Vitamin M / Vitamin B9)	4	4	7	10
C4	Vitamin B1 (Thiamin)	1	1	7	10
C5	L-Carnitine (Vitamin BT)	1	0	6	2
C6	Poly-N-vinylpyrrolidone	1	0	4	Nil
C7	Inulin	11	0	12	10
C8	Poly Cysteine	1	1	1	7
C9	Chitosan	1	1	1	10
C10	Pectin	3	0	6	10
C11	Poly(propylene glycol)	1	0	4	4
C12	Poly(propylene imine)	0	1	3	Nil
C13	Poly (lactic-co-glycolic acid)	2	0	2	8
C14	Deoxycholic acid	2	0	14	7

Table 3

Physico-chemical properties of drug delivering molecules

S. No	Carriers	ADME Descriptors					
		Aqueous Solubility	Blood Brain barrier Penetration	CYP 2D6 Binding	Hepatotoxicity	Intestinal Absorption	Plasma Protein Binding
1	C1	-9.057	4.0	0.277	0.509	3.0	0.0
2	C2	-1.432	-1.229	0.069	0.264	0.0	0.0
3	C3	-3.378	4.0	0.277	0.662	3.0	0.0
4	C4	-1.335	-1.253	0.336	0.423	0.0	0.0
5	C5	2.734	4.0	0.059	0.066	3.0	0.0
6	C6	-0.550	4.0	0.059	0.324	1.0	0.0
7	C7	0.894	4.0	0.108	0.0139	3.0	0.0
8	C8	0.335	-1.362	0.059	0.033	0.0	0.0
9	C9	1.736	4.0	0.029	0.013	3.0	0.0
10	C10	1.595	4.0	0.029	0.059	3.0	0.0
11	C11	0.907	-1.039	0.0590	0.059	0.0	0.0
12	C12	-0.097	4.0	0.059	0.324	1.0	0.0
13	C13	0.768	-1.713	0.059	0.006	0.0	0.0
14	C14	-4.409	-0.154	0.485	0.026	0.0	1.0

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 4
Toxicity studies for Drug delivering molecules by Discovery Studio and Toxtree

Toxicity Prediction by Discovery Studio									Toxicity Prediction by Toxtree
S. No	Carriers	FDA Rodent Carcinogenicity	Mutagenecity	Rat oral LD50 (g/Kg Body weight)	Rat Maximum tolerated dose (g/Kg Body weight)	Skin Irritancy	Skin sensitization	Aerobic Biodegradability	Toxicity
1	C1	Non-carcinogen	Non-mutagen	0.093	0.000	Mild-Irritant	Non-sensitizer	Degradable	High
2	C2	Non-carcinogen	Non-mutagen	1.109	0.193	Non-Irritant	Non-sensitizer	Degradable	Low
3	C3	Non-carcinogen	Non-mutagen	2.819	1.391	Non-Irritant	Weak-sensitizer	Non-Degradable	High
4	C4	carcinogen	Non-mutagen	1.308	0.097	Non-Irritant	Strong-sensitizer	Non-Degradable	Low
5	C5	Non-carcinogen	Non-mutagen	1.101	0.175	Mild-Irritant	Non-sensitizer	Degradable	Low
6	C6	carcinogen	Non-mutagen	1.634	0.181	Mild-Irritant	Non-sensitizer	Degradable	High
7	C7	Non-carcinogen	Non-mutagen	20.789	0.000	Mild-Irritant	Non-sensitizer	Degradable	Low
8	C8	Non-carcinogen	Non-mutagen	0.514	0.653	Non-Irritant	Non-sensitizer	Degradable	Low
9	C9	Non-carcinogen	Non-mutagen	3.241	0.268	Mild-Irritant	Non-sensitizer	Degradable	High
10	C10	Non-carcinogen	Non-mutagen	3.576	0.525	Non-Irritant	Non-sensitizer	Degradable	Low
11	C11	Non-carcinogen	Non-mutagen	12.098	0.187	Mild-Irritant	Non-sensitizer	Degradable	High
12	C12	Non-carcinogen	Non-mutagen	0.055	0.089	Mild-Irritant	Non-sensitizer	Non-Degradable	High
13	C13	Non-carcinogen	Non-mutagen	2.982	0.427	Non-Irritant	Non-sensitizer	Degradable	Low
14	C14	Non-carcinogen	Non-mutagen	6.358	0.190207	Severe-Irritant	Weak-sensitizer	Degradable	High

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 5
Conjugation results of Human Insulin Monomer (PDB ID: 1HLS), with all listed drug delivering molecules individually

S.No	Conjugate	Binding site (LYS B29)	LIP DOCK SCORE -	Binding site (PHE B1)	LIP DOCK SCORE -	Other Binding site	LIP DOCK SCORE
C1	Vitamin B12	-	-	-	-	-	-
C2	Vitamin H	-	-	-	-	CYS A11 HIS B10 GLN B4	86.2835
C3	Vitamin M	-	-	-	-	-	-
C4	Vitamin B1	-	-	-	-	CYS A6 CYS A11 LEU B6	94.1144
C5	Vitamin BT	-	-	-	-	LEU A13 VAL B2	68.0592
C6	poly-N-vinylpyrrolidone	-	-	-	-	-	-
C7	Inulin	-	-	PHE B1	117.663	-	-
C8	Poly Cysteine	-	-	-	-	GLN B4	53.1597
C9	Chitosan	-	-	-	-	CYS A11, VAL B2, ASN B3	74.4251
C10	PECTIN	-	-	-	-	CYS A11	63.0902
C11	POLY (PROPYLENE GLYCOL)	-	-	-	-	VAL B2	76.0238
C12	poly(propylene imine)	-	-	-	-	-	-
C13	Poly (lactic-co-glycolic acid)	-	-	PHE B1	68.6737	CYS A11	68.6737
C14	Deoxycholic acid	-	-	-	-	-	-

Table 6
Conjugation results of Human insulin hexamer (PDB ID: 1AIO), with all listed drug delivering molecules individually

S.No	Conjugate	Binding site (LYS B29)	LIP DOCK SCORE -	Binding site (PHE B1)	LIP DOCK SCORE -	Other Binding site	LIP DOCK SCORE
C1	Vit B12	-	-	-	-	-	-
C2	Vit H	-	-	-	-	GLN B4 ARG B22	103.231
C3	Vit M	-	-	-	-	GLY A1 THR B27	114.324
C4	Vit B1	LYS B29	79.8834	-	-	-	-
C5	Vit BT	-	-	-	-	ARG B22	84.7767
C6	poly-N-vinylpyrrolidone	-	-	-	-	-	-
C7	Inulin	LYS B29	94.3543	-	-	-	-
C8	Poly Cysteine	-	-	-	-	TYR B16 GLU B21	58.5629
C9	Chitosan	-	-	-	-	TYR B16 TYR B26	90.016
C10	PECTIN	-	-	-	-	GLN B4	91.0549
C11	POLY (PROPYLENE GLYCOL)	-	-	-	-	GLU B21 ARG B22	71.5126
C12	poly(propylene imine)	-	-	-	-	-	-
C13	Poly (lactic-co-glycolic acid)	-	-	-	-	GLY B20 ARG B22	82.2602
C14	Deoxycholic acid	-	-	-	-	GLY A1 ILE A2	60.763

Table 7
Conjugation results of Insulin Lispro (PDB ID: 1 LPH), with all listed drug delivering molecules individually

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

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

S. No	Conjugates	Interaction with leucine-rich repeat domain (L1, residues 1-157)	Interaction with C-terminus of the α-chain (αCT, residues 704-715)
1	Insulin Monomer (1HLS)- DDM Conjugates	No Interaction	No Interaction
2	Insulin Hexamer (1AIO)- DDM Conjugates	ARG86 ASN34	No Interaction
3	Insulin Lispro (1LPH) - DDM Conjugates	ARG86 ASN90 ARG114	No Interaction

Supplementary figure 1

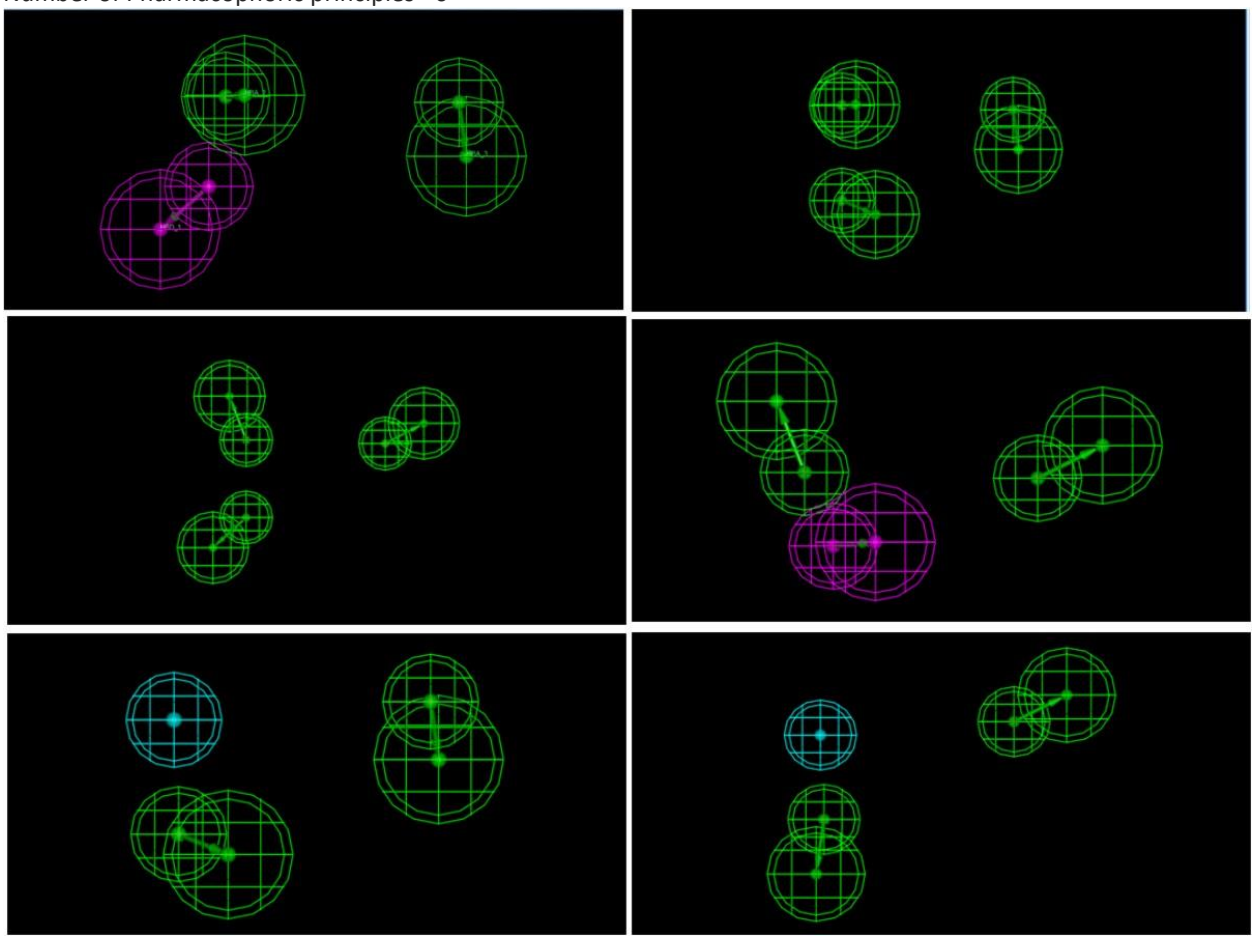
Pharmacophoric principles of Drug delivering molecules are illustrated by Discovery Studio software. Acceptors (Green in color), Donors (Magenta in color) of Pharmacophoric principles of all Drug delivering molecules are demonstrated and differentiated by color. C1 - Vitamin B12 (cobalamin); C2 - Vitamin 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

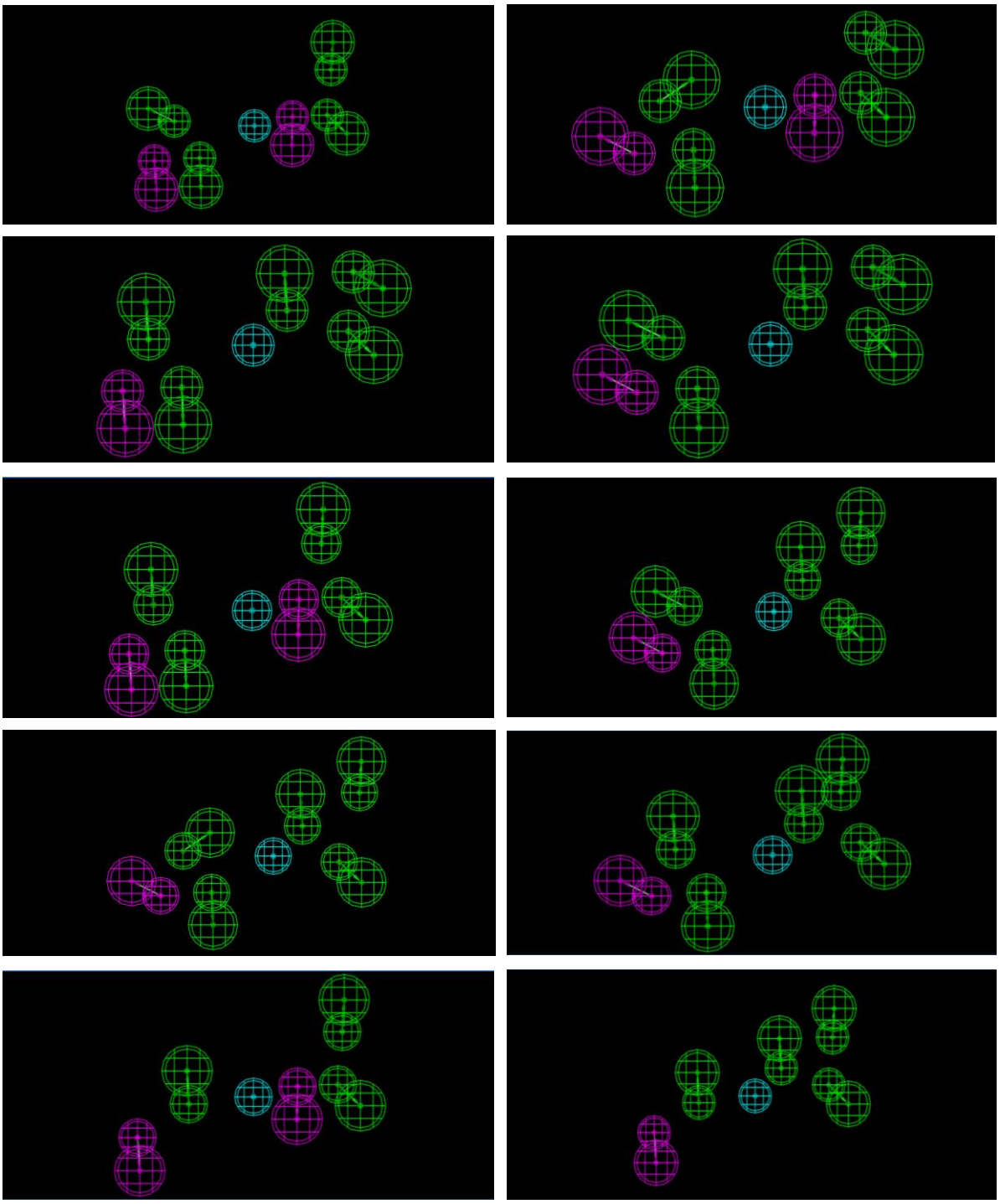
Supplementary Figure (1a): C1 - Vitamin B12 (cobalamin)

Number of Pharmacophoric principles – Nil

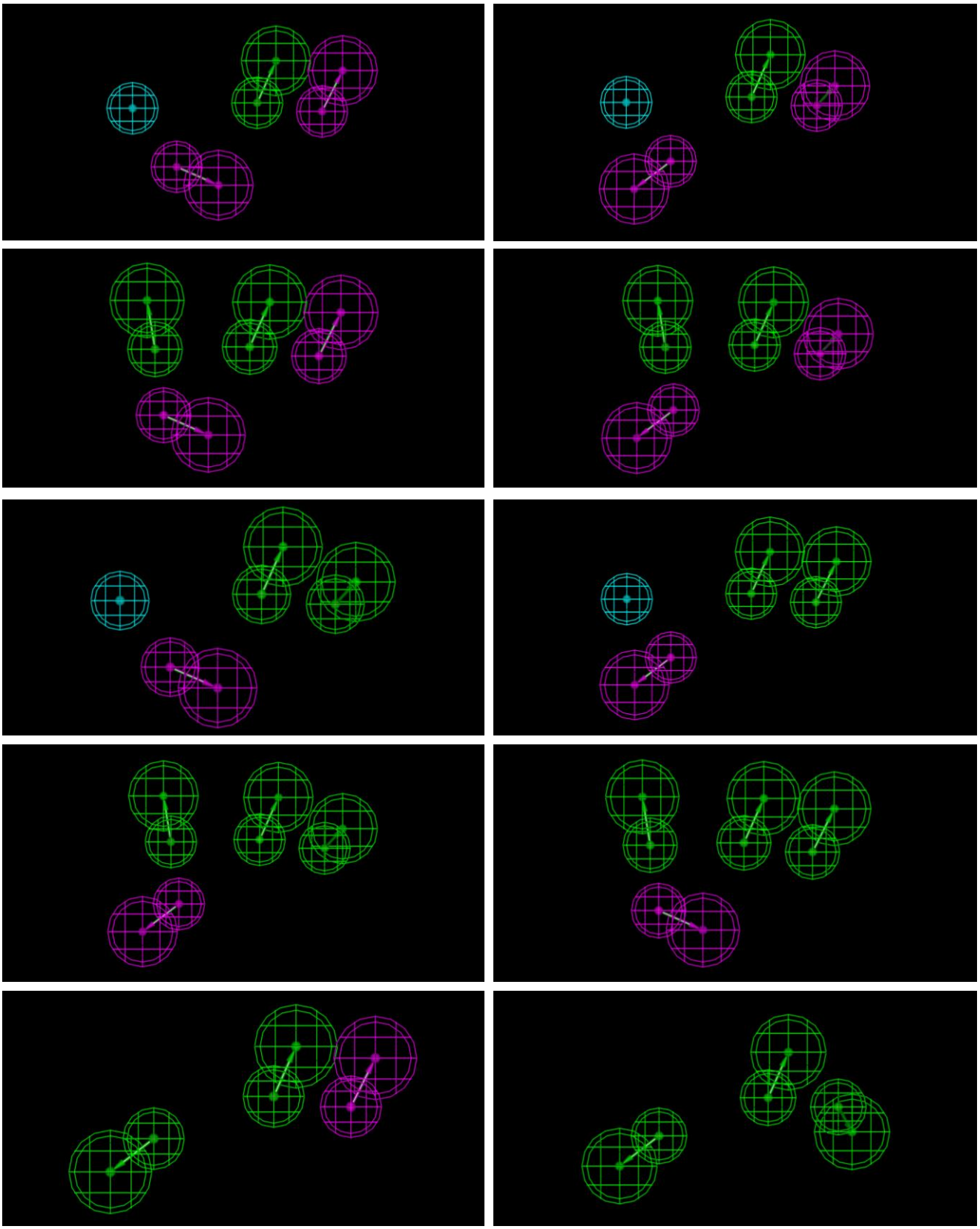
Supplementary Figure (1b): C2 - Vitamin H (Biotin)

Number of Pharmacophoric principles –6



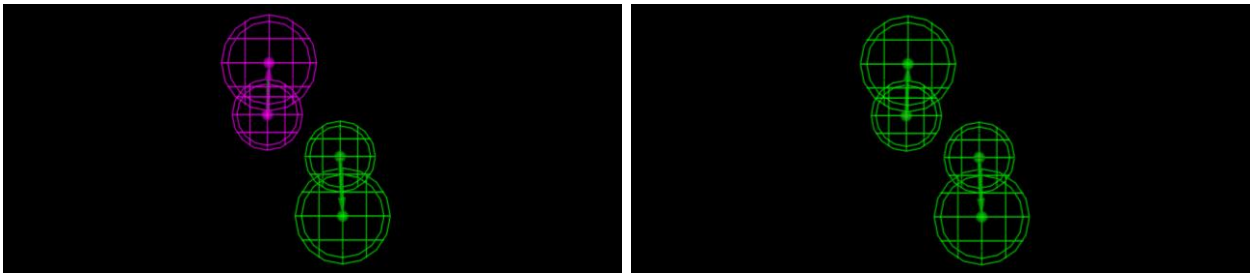


Supplementary Figure (1c): C3 - Folic acid (Vitamin M / Vitamin B9)
Number of Pharmacophoric principles –10



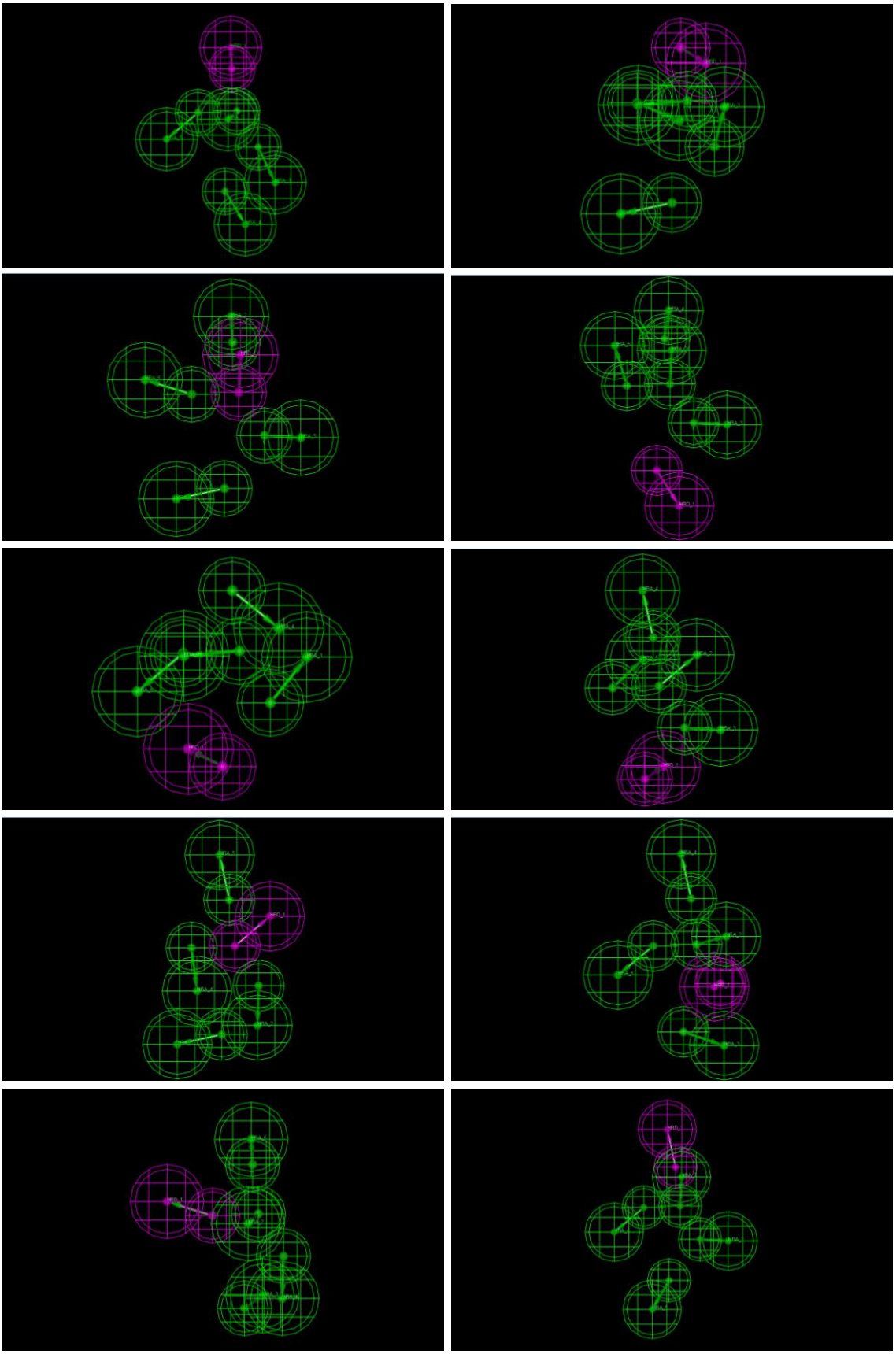
Supplementary Figure (1d): C4 - Vitamin B1 (Thiamin)
Number of Pharmacophoric principles –10

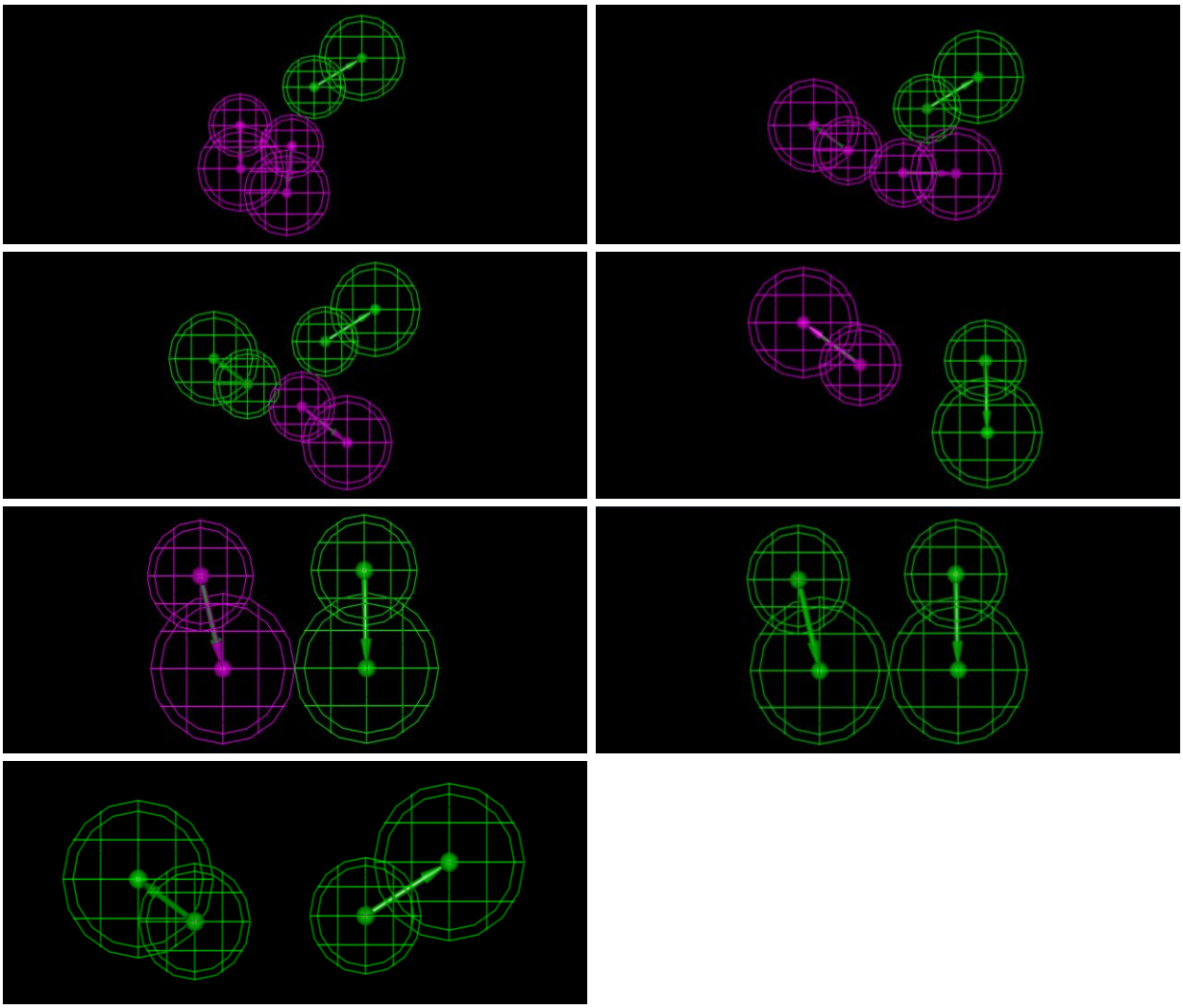
Supplementary Figure (1e): C5 - L-Carnitine (Vitamin BT)
Number of Pharmacophoric principles –2



Supplementary Figure (1f): C6- Poly-N-vinylpyrrolidone
Number of Pharmacophoric principles –Nil

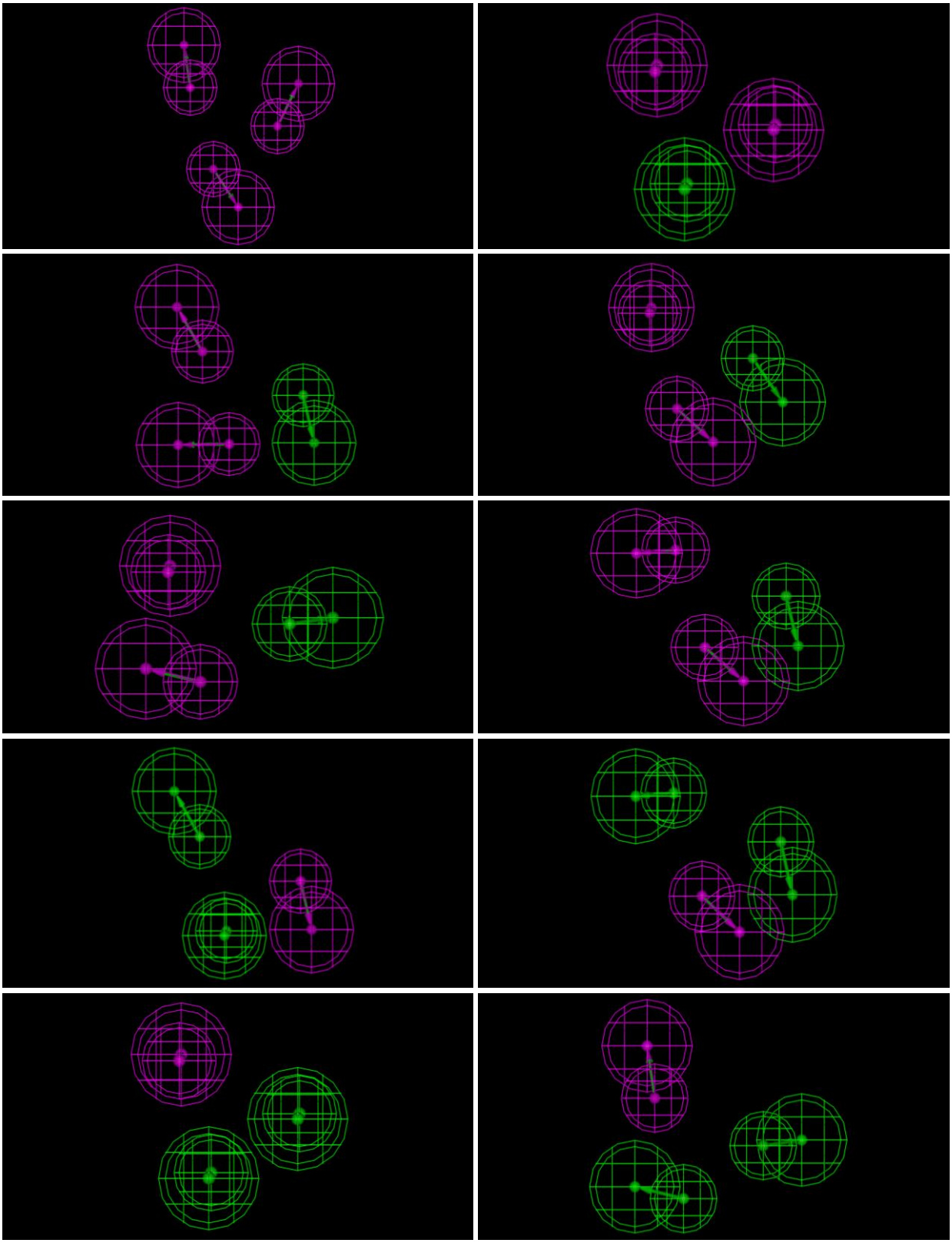
Supplementary Figure (1g): C7- Inulin
Number of Pharmacophoric principles –10



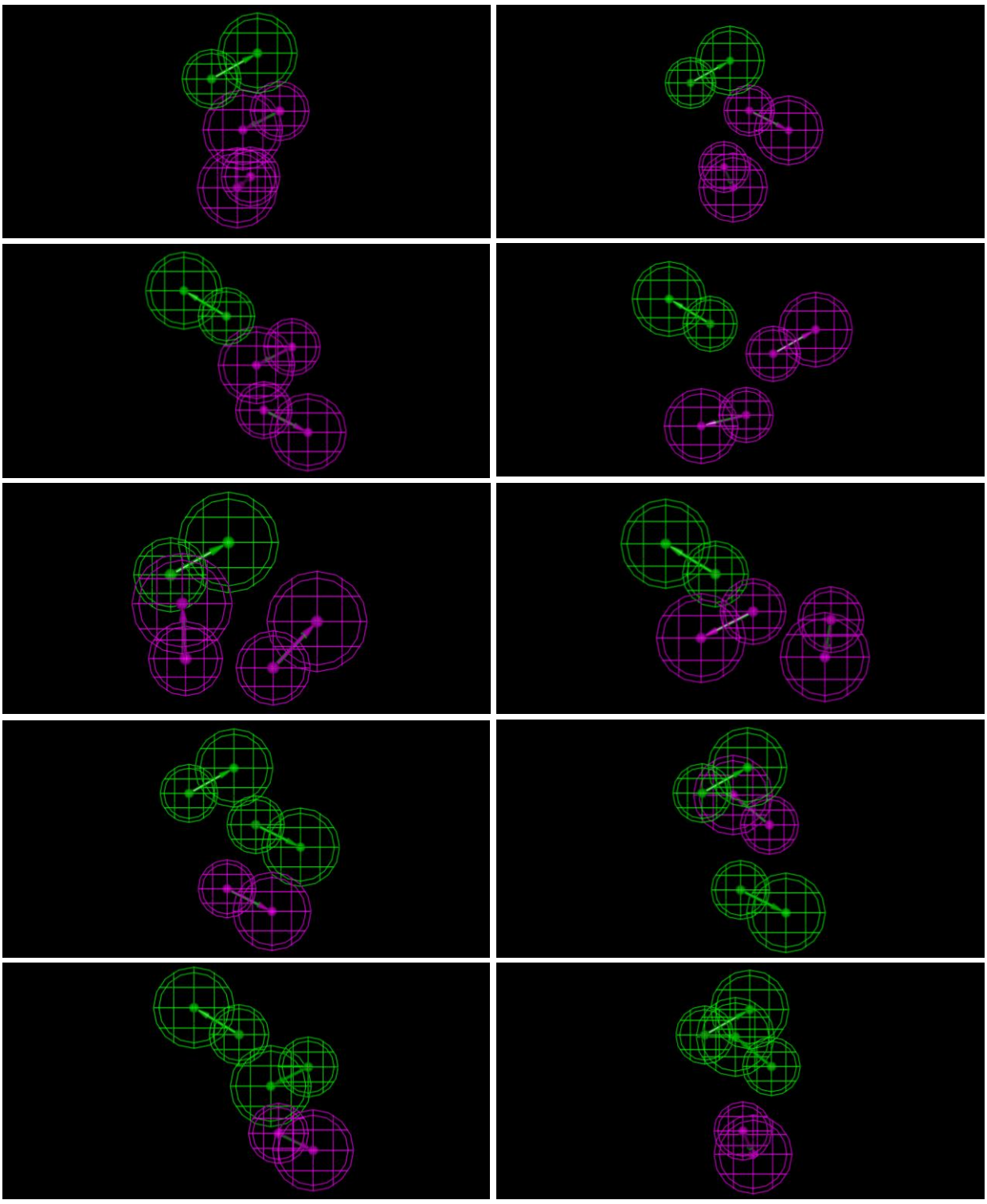


Supplementary Figure (1h): C8- Poly Cysteine
Number of Pharmacophoric principles –7

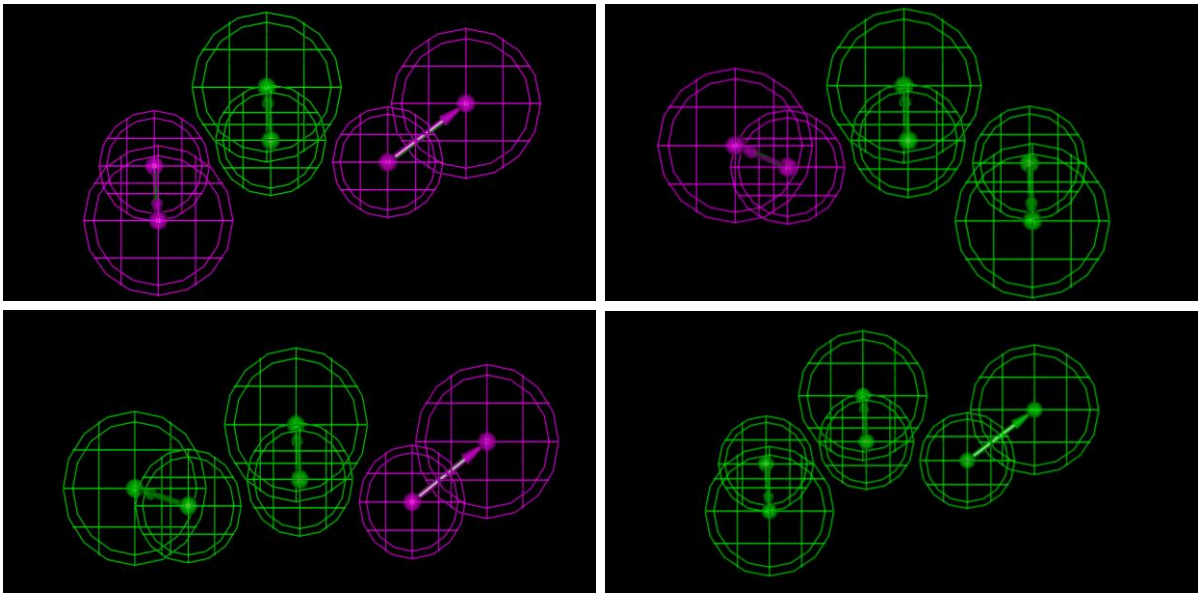
Supplementary Figure (1i): C9 – Chitosan
Number of Pharmacophoric principles –10



Supplementary Figure (1j): C10 – Pectin
Number of Pharmacophoric principles –10

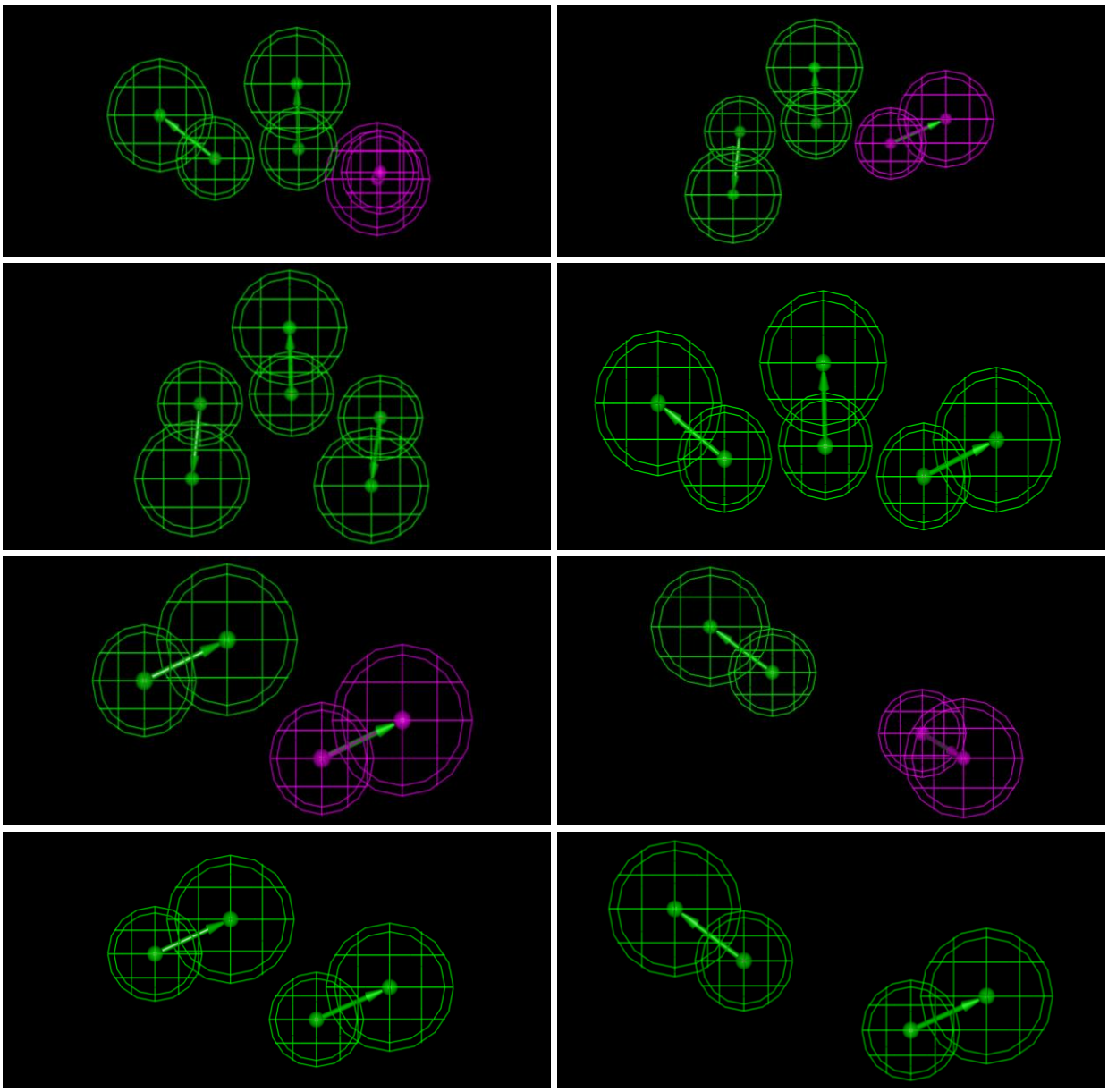


Supplementary Figure (1k): C11 – Poly(propylene glycol)
Number of Pharmacophoric principles –4

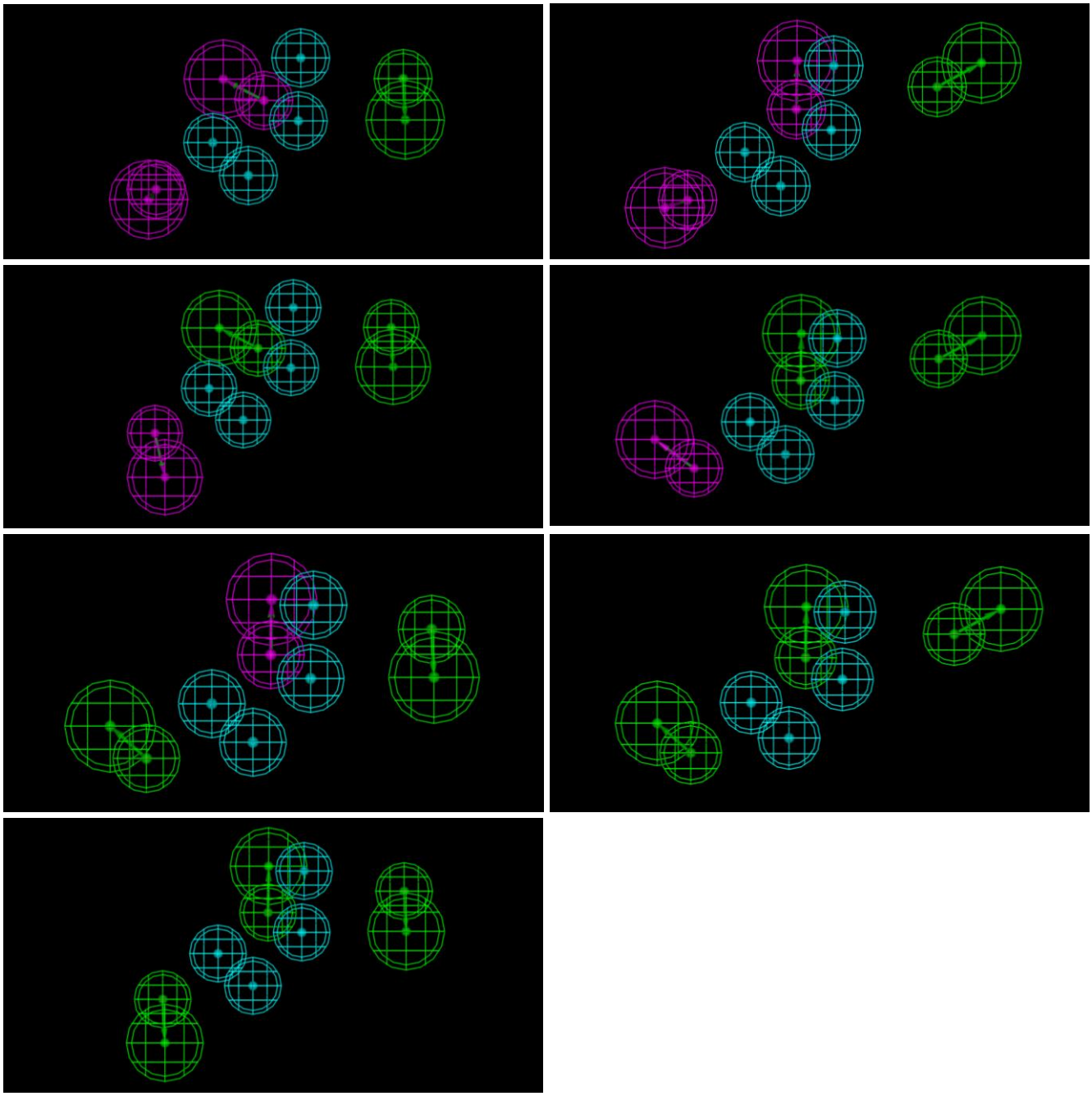


Supplementary Figure (1L): C12 – Poly(propylene imine)
Number of Pharmacophoric principles –Nil

Supplementary Figure (1m): C13 – Poly (lactic-co-glycolic acid)
Number of Pharmacophoric principles –8



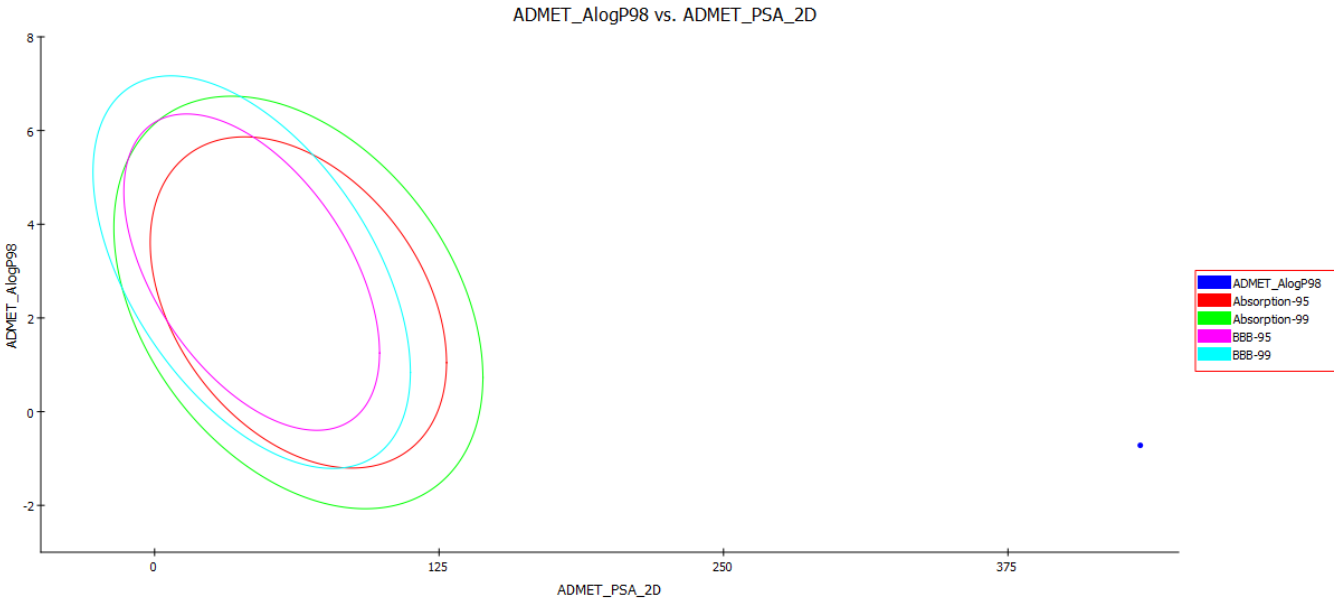
Supplementary Figure (1n): C14 – Deoxycholic acid
Number of Pharmacophoric principles –7



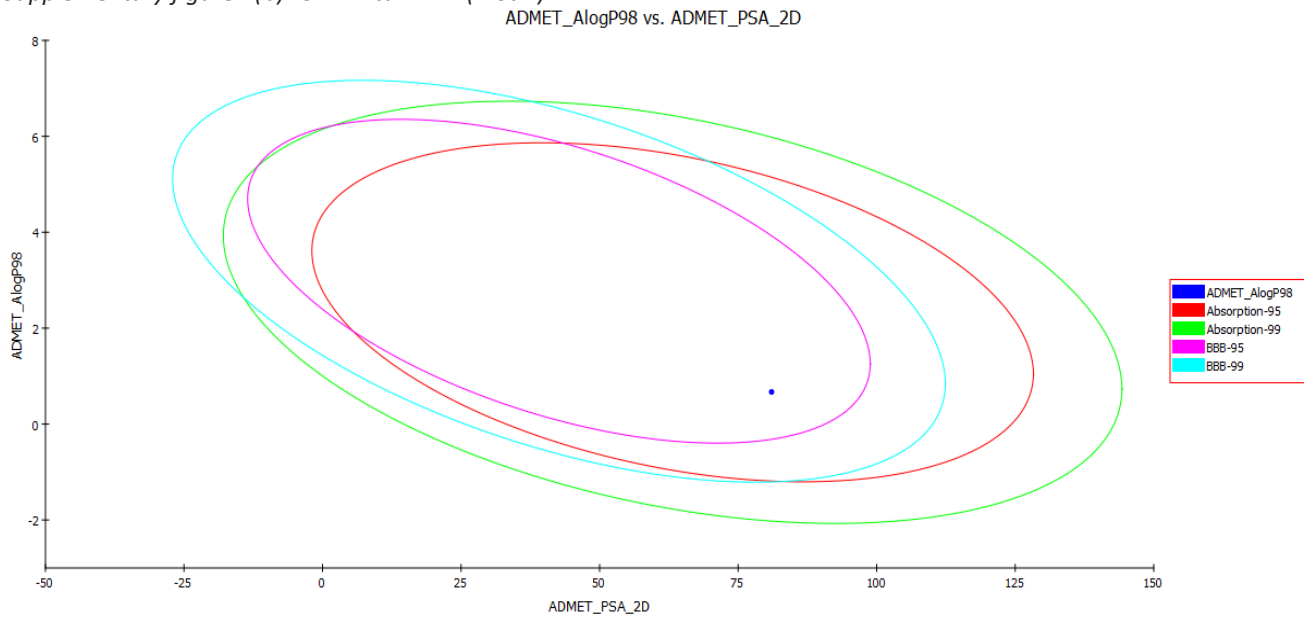
Supplementary figure 2

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

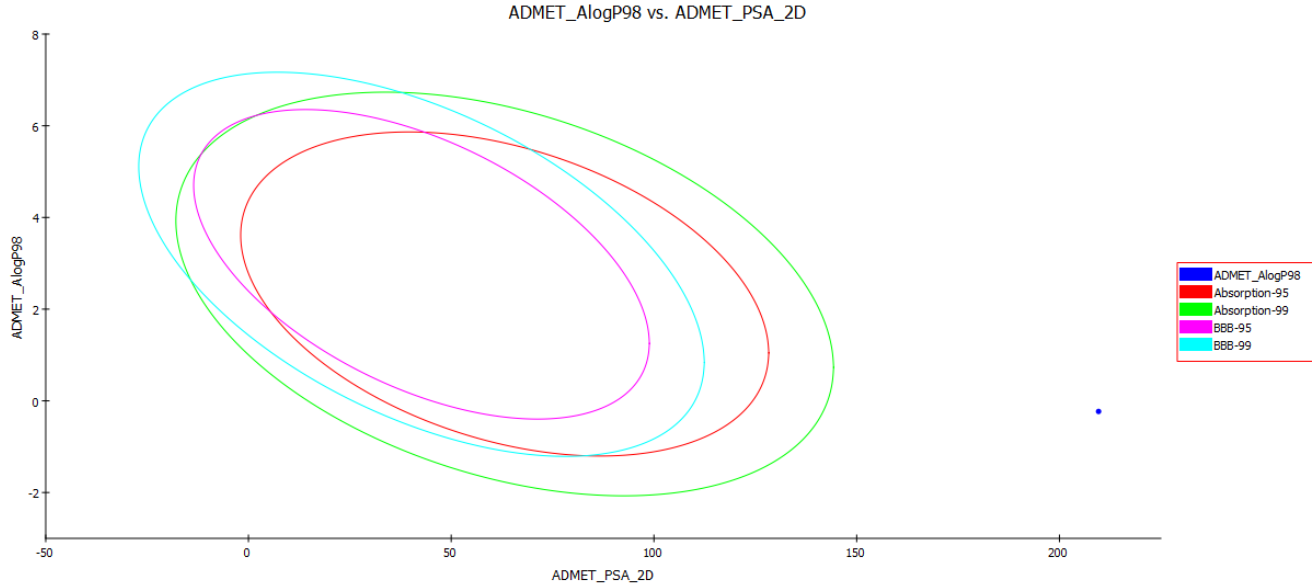
Supplementary figure 2(a)- C1 - Vitamin B12



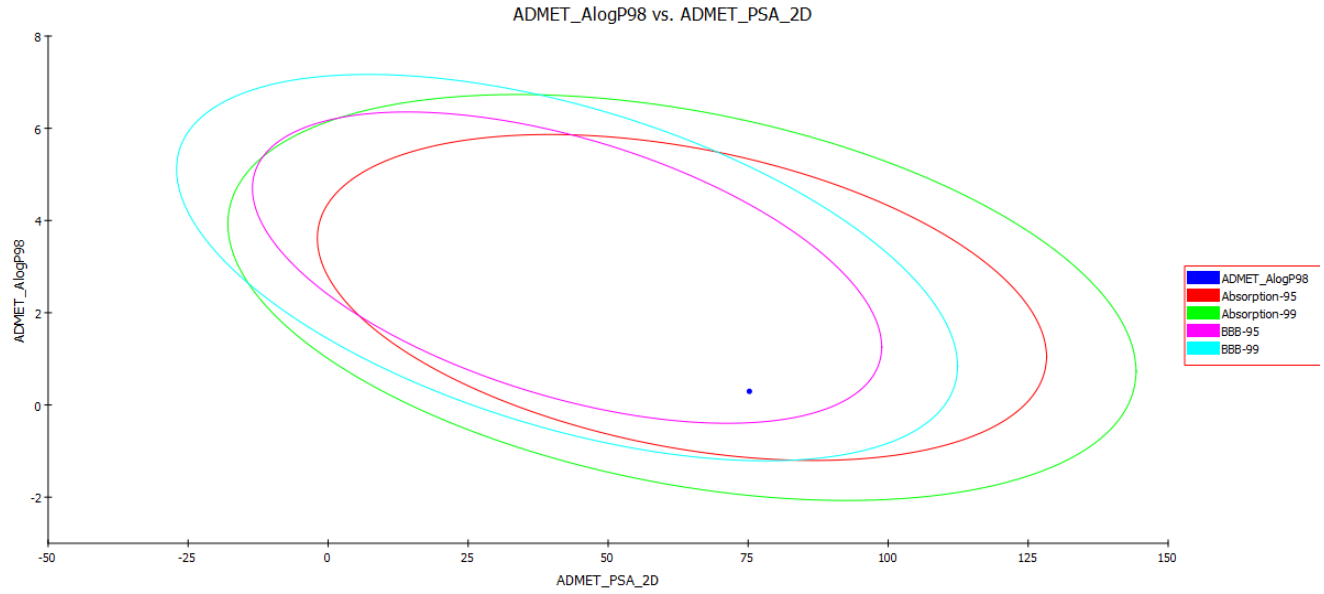
Supplementary figure 2(b)- C2 - Vitamin H (Biotin)



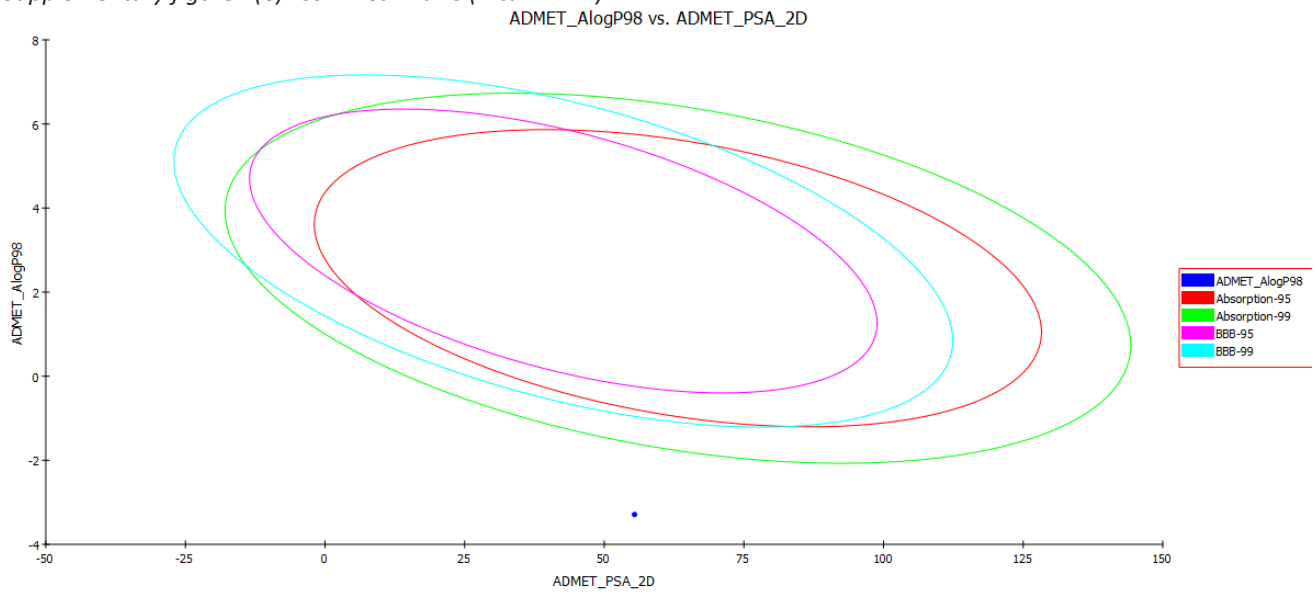
Supplementary figure 2(c)- C3 - Folic acid (Vitamin M / Vitamin B9)



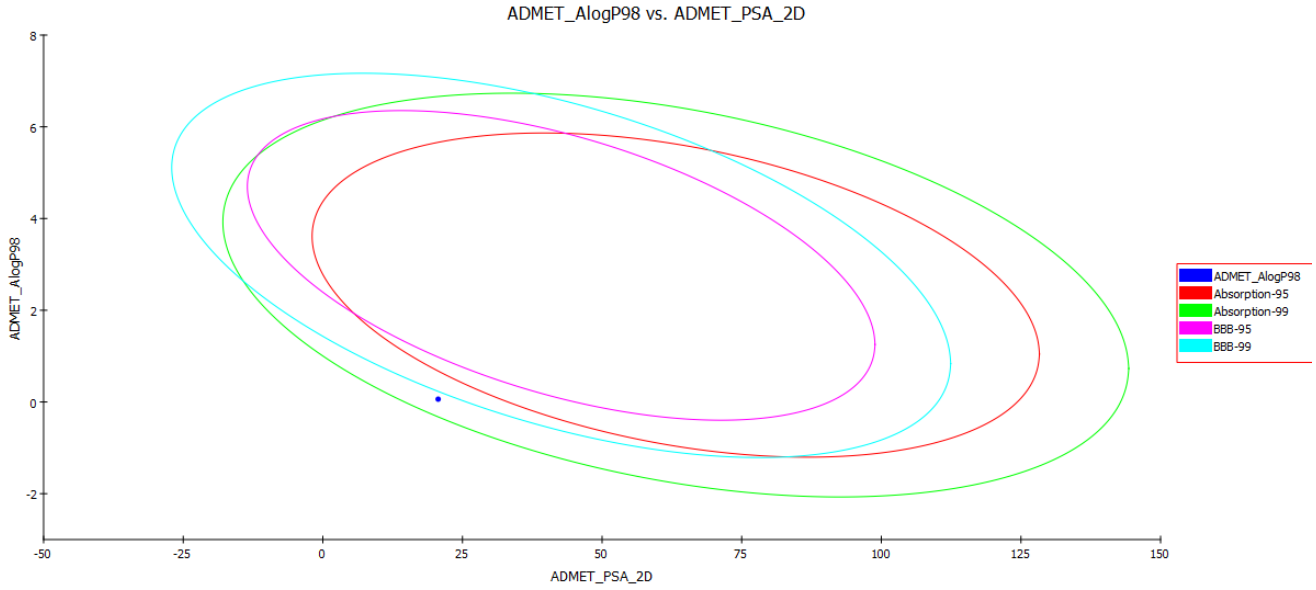
Supplementary figure 2(d)- C4 - Vitamin B1 (Thiamin)



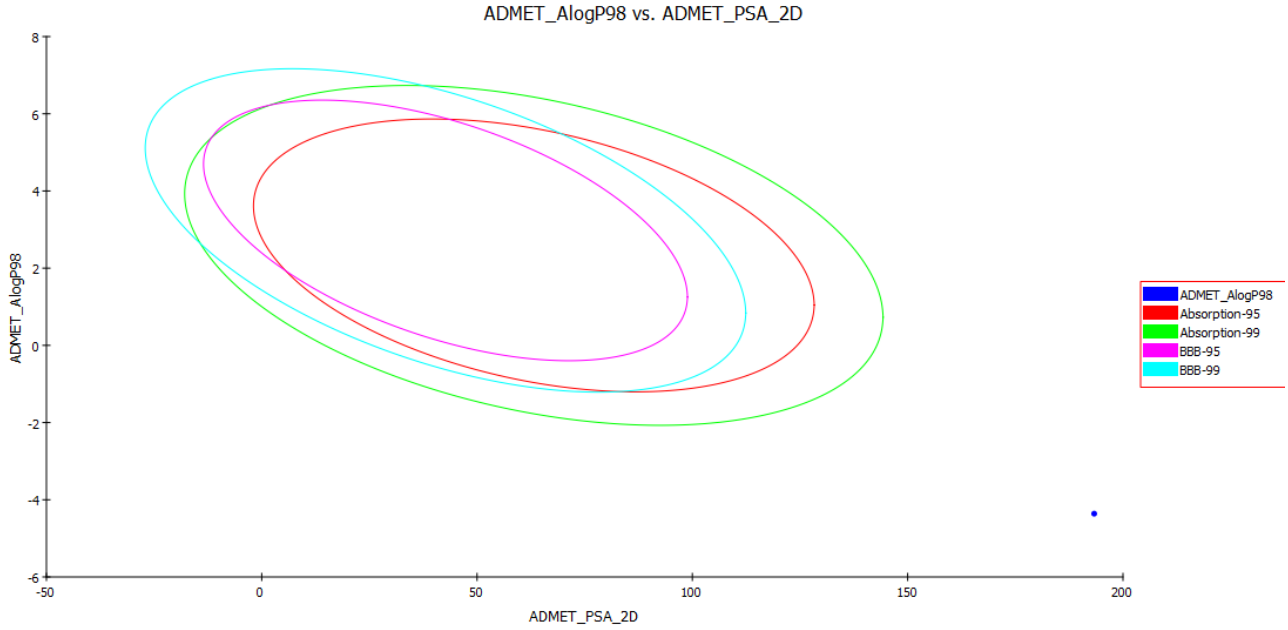
Supplementary figure 2(e)- C5 - L-Carnitine (Vitamin BT)



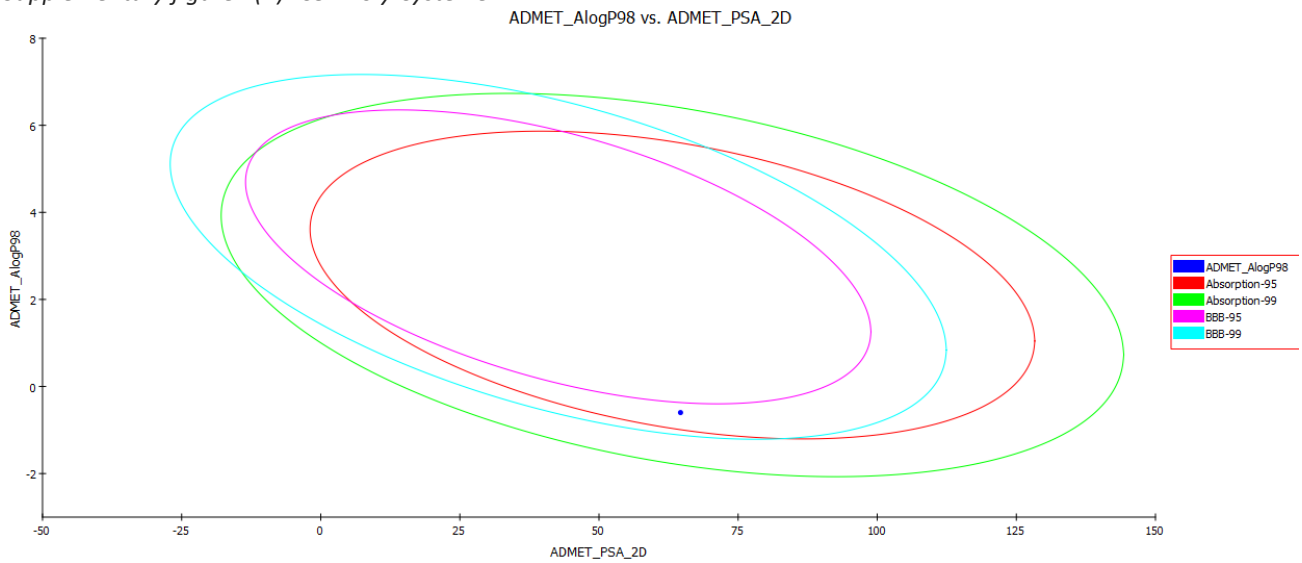
Supplementary figure 2(f)- C6 - Poly-N-vinylpyrrolidone



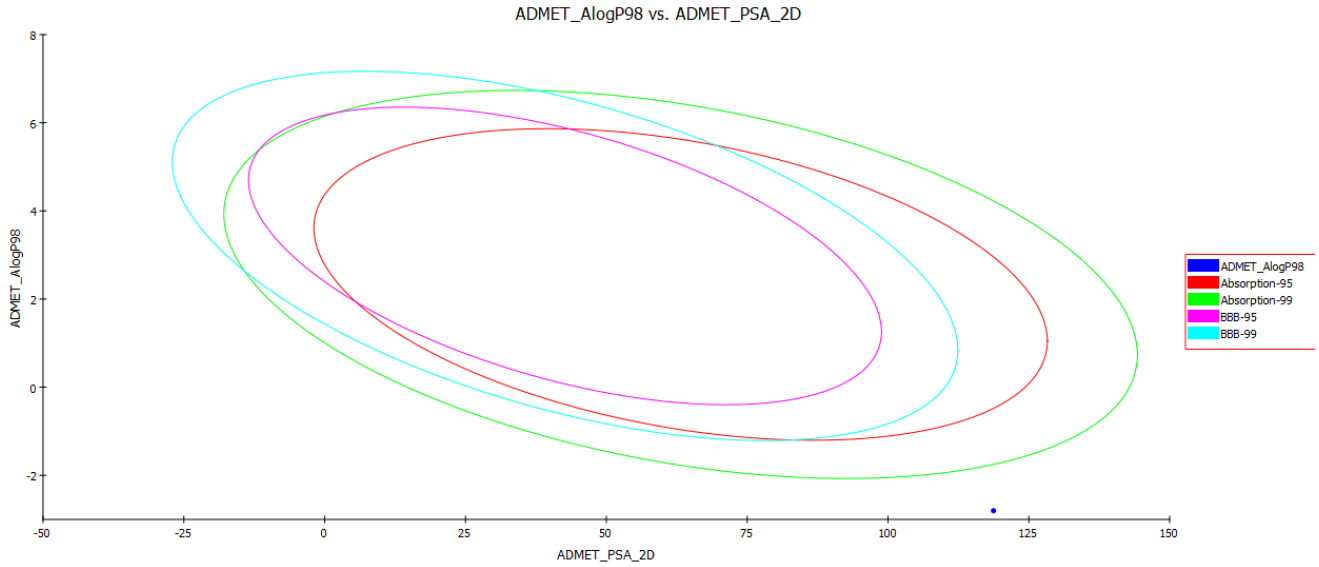
Supplementary figure 2(g)- C7- Inulin



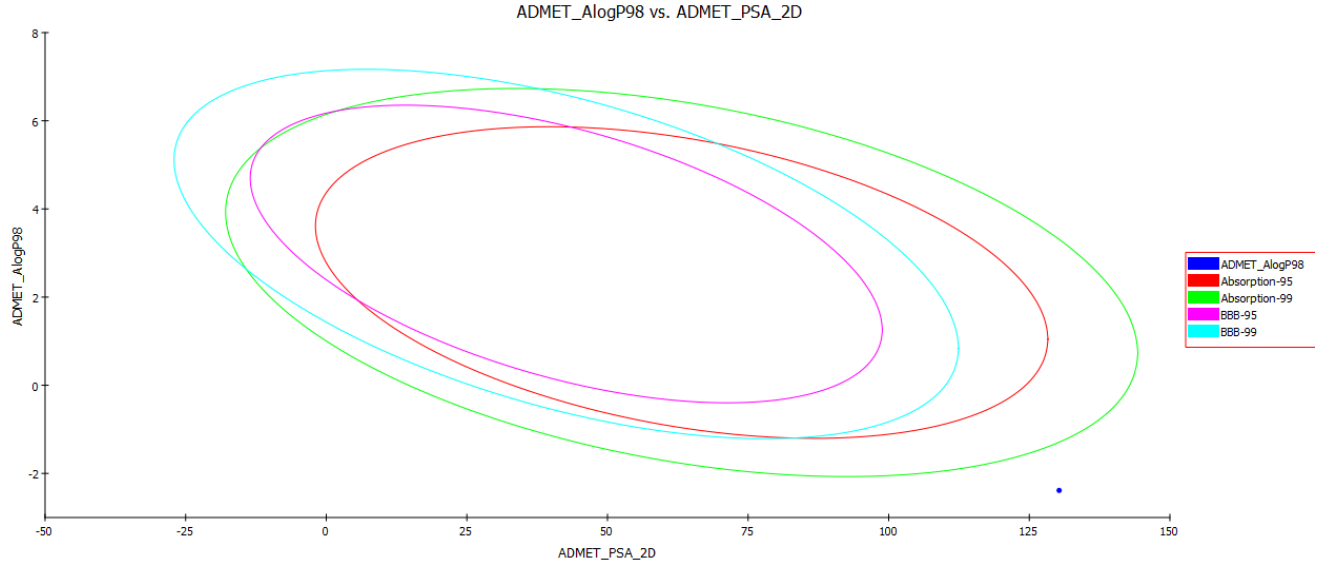
Supplementary figure 2(h)- C8 - Poly Cysteine



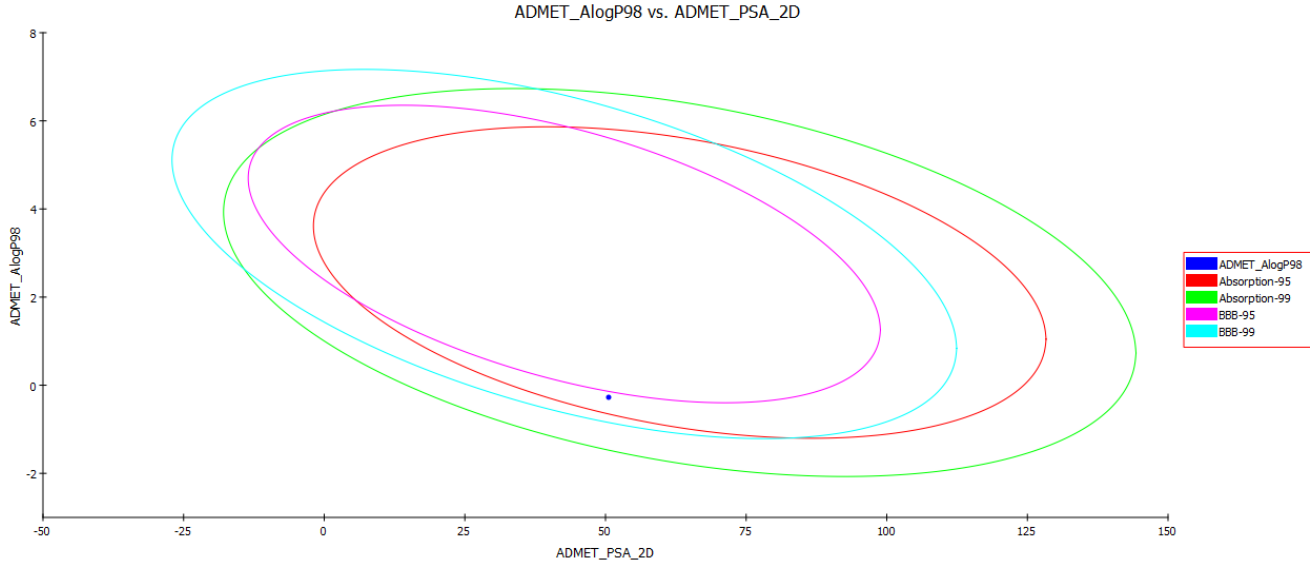
Supplementary figure 2(i)- C9 – Chitosan



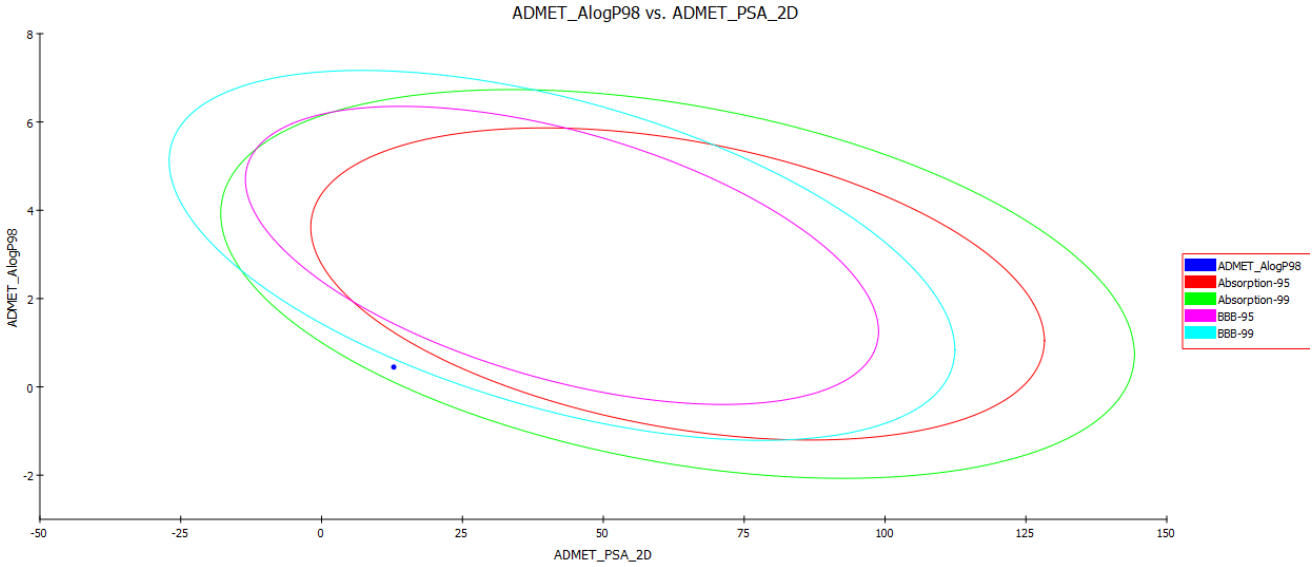
Supplementary figure 2(j)- C10 – Pectin



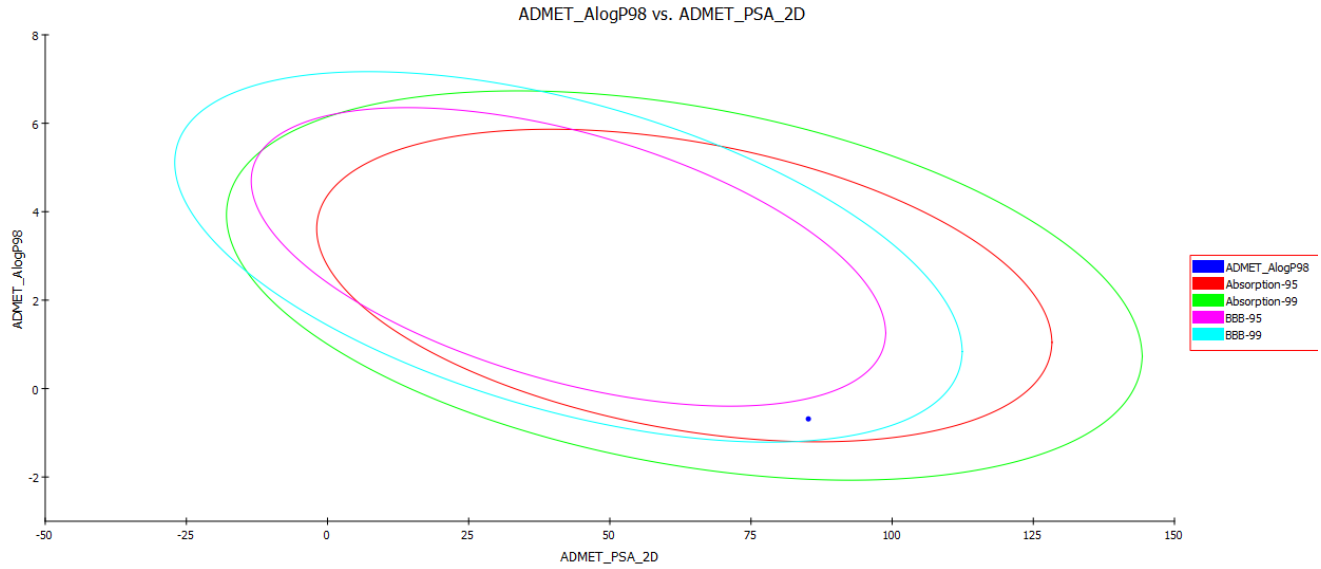
Supplementary figure 2(k)- C11 - Poly(propylene glycol)



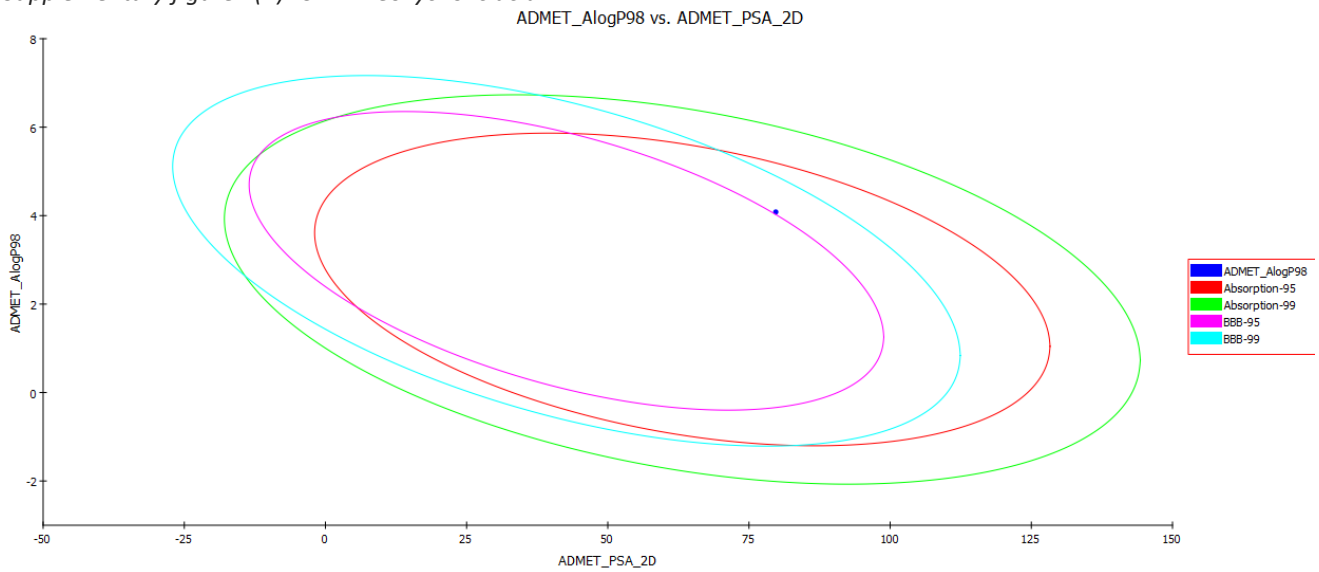
Supplementary figure 2(l)- C12 - Poly(propylene imine)



Supplementary figure 2(m)- C13 - Poly (lactic-co-glycolic acid)



Supplementary figure 2(n)- C14 - Deoxycholic acid

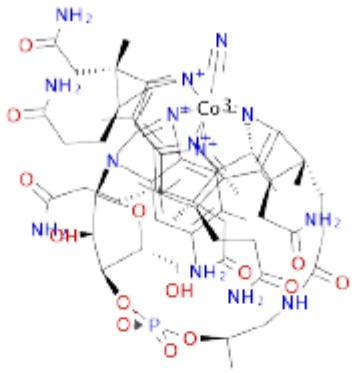


Supplementary figure 3

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)

FDA Rodent Carcinogenicity



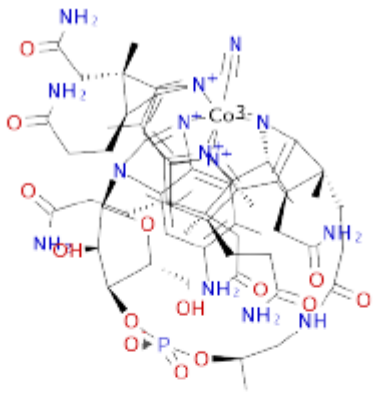
$C_{63}H_{88}CoN_{14}O_{14}P$
Molecular Weight: 1355.36518
ALogP: -0.442
Rotatable Bonds: 16
Acceptors: 16
Donors: 9

Model Prediction

Prediction: 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 esimated 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



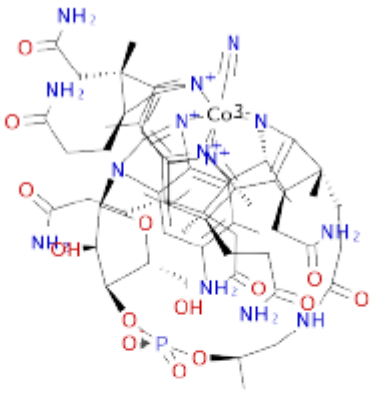
$C_{63}H_{88}CoN_{14}O_{14}P$
Molecular Weight: 1355.36518
ALogP: -0.442
Rotatable Bonds: 16
Acceptors: 16
Donors: 9

Model Prediction

Prediction: 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 esimated 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.

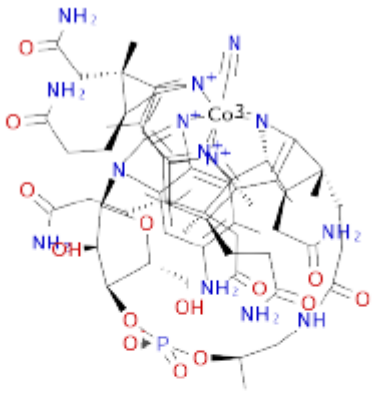
Rat oral LD50 (g/Kg Body weight)



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

Model Prediction
Prediction: 0.093
Unit: g/kg_body_weight

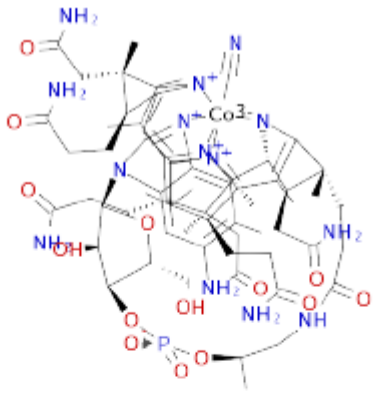
Rat Maximum tolerated dose (g/Kg Body weight)



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

Model Prediction
Prediction: 0.000
Unit: g/kg_body_weight

Skin Irritancy



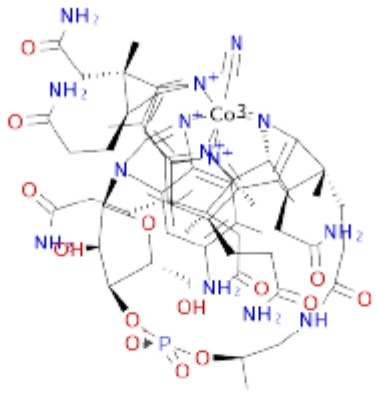
$C_{63}H_{88}CoN_{14}O_{14}P$
Molecular Weight: 1355.36518
ALogP: -0.442
Rotatable Bonds: 16
Acceptors: 16
Donors: 9

Model Prediction

Prediction: 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 esimated 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



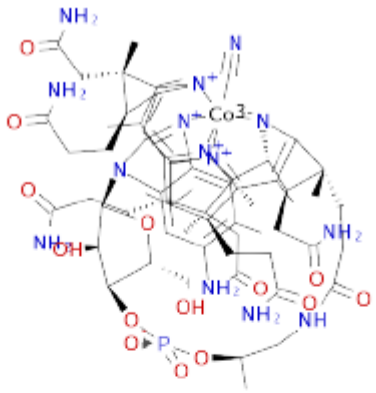
$C_{63}H_{88}CoN_{14}O_{14}P$
Molecular Weight: 1355.36518
ALogP: -0.442
Rotatable Bonds: 16
Acceptors: 16
Donors: 9

Model Prediction

Prediction: 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 esimated 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



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

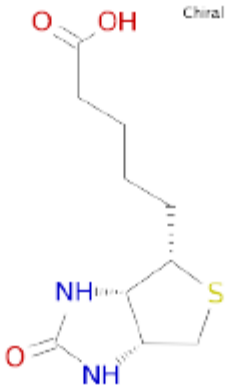
Model Prediction

Prediction: 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 esimated 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



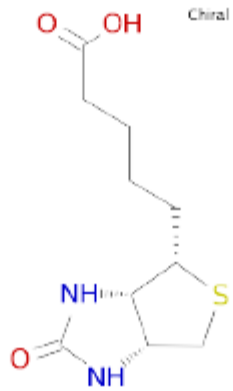
$C_{10}H_{16}N_2O_3S$
Molecular Weight: 244.31064
ALogP: 0.67
Rotatable Bonds: 5
Acceptors: 4
Donors: 3

Model Prediction

Prediction: Non-Carcinogen
Probability: 0.435
Enrichment: 0.845
Bayesian Score: -2.888

Prediction: Positive if the Bayesian score is above the estimated best cutoff value.
Probability: The esimated 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



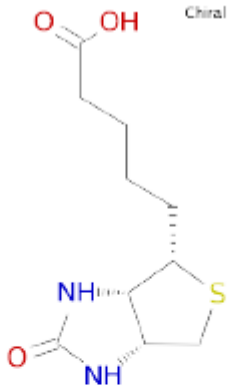
$C_{10}H_{16}N_2O_3S$
Molecular Weight: 244.31064
ALogP: 0.67
Rotatable Bonds: 5
Acceptors: 4
Donors: 3

Model Prediction

Prediction: Non-Mutagen
Probability: 0.002
Enrichment: 0.004
Bayesian Score: -29.037

Prediction: Positive if the Bayesian score is above the estimated best cutoff value.
Probability: The esimated 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.

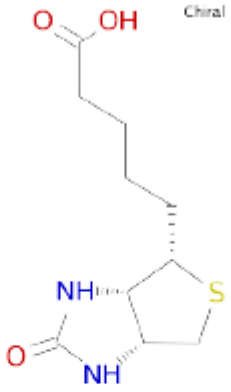
Rat oral LD50 (g/Kg Body weight)



$C_{10}H_{16}N_2O_3S$
Molecular Weight: 244.31064
ALogP: 0.67
Rotatable Bonds: 5
Acceptors: 4
Donors: 3

Model Prediction
Prediction: 1.109
Unit: g/kg_body_weight

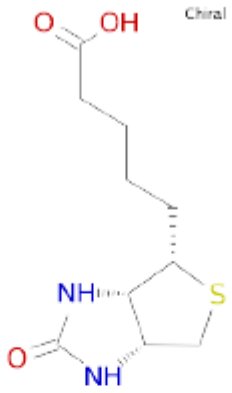
Rat Maximum tolerated dose (g/Kg Body weight)



$C_{10}H_{16}N_2O_3S$
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



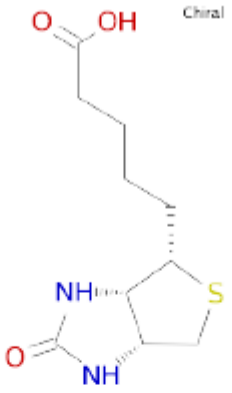
$C_{10}H_{16}N_2O_3S$
Molecular Weight: 244.31064
ALogP: 0.67
Rotatable Bonds: 5
Acceptors: 4
Donors: 3

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 esimated 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

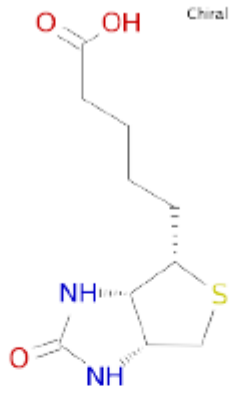


$C_{10}H_{16}N_2O_3S$
Molecular Weight: 244.31064
ALogP: 0.67
Rotatable Bonds: 5
Acceptors: 4
Donors: 3

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 esimated 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.



$C_{10}H_{16}N_2O_3S$
Molecular Weight: 244.31064
ALogP: 0.67
Rotatable Bonds: 5
Acceptors: 4
Donors: 3

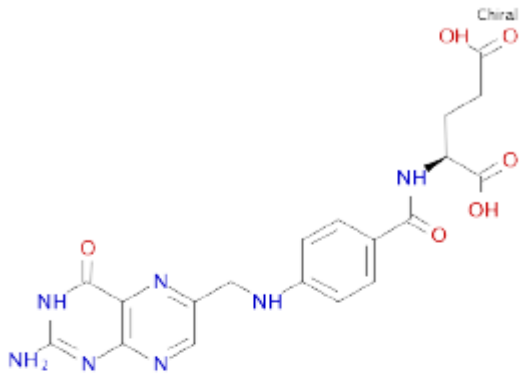
Model Prediction

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 esimated 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)

FDA Rodent Carcinogenicity



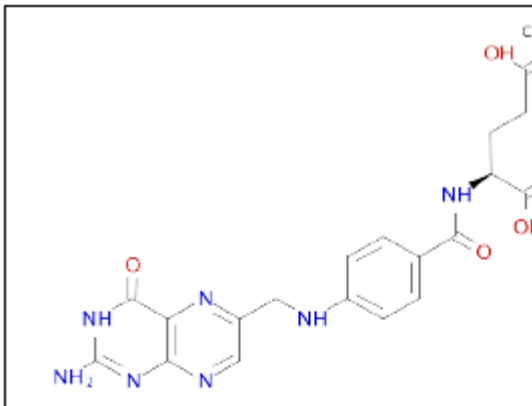
$C_{19}H_{19}N_7O_6$
Molecular Weight: 441.39745
ALogP: -0.232
Rotatable Bonds: 9
Acceptors: 11
Donors: 6

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 esimated 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



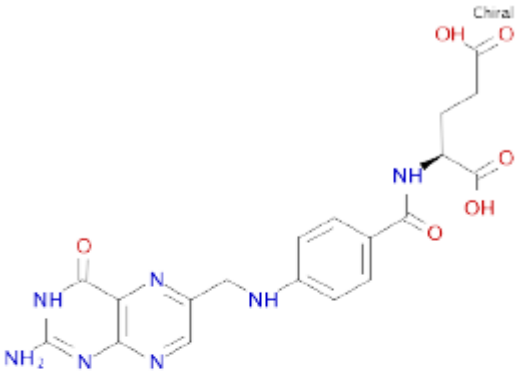
$C_{19}H_{19}N_7O_6$
Molecular Weight: 441.39745
ALogP: -0.232
Rotatable Bonds: 9
Acceptors: 11
Donors: 6

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 esimated 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.

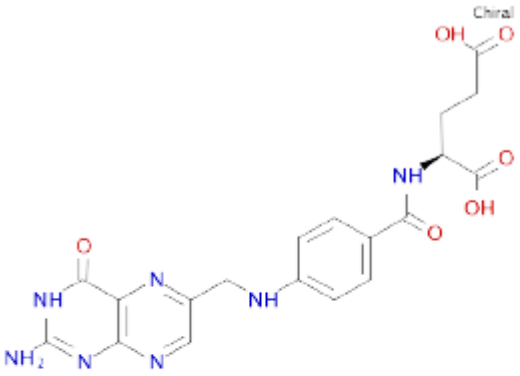
Rat oral LD50 (g/Kg Body weight)



$C_{19}H_{19}N_7O_6$
Molecular Weight: 441.39745
ALogP: -0.232
Rotatable Bonds: 9
Acceptors: 11
Donors: 6

Model Prediction
Prediction: 2.819
Unit: g/kg_body_weight

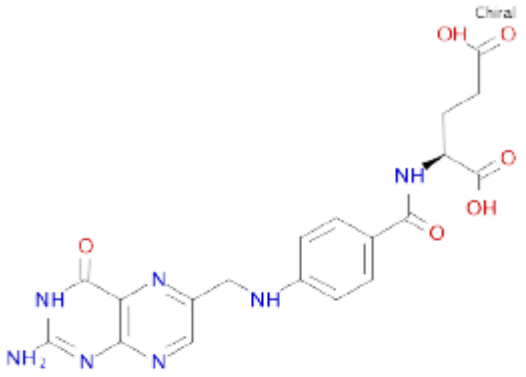
Rat Maximum tolerated dose (g/Kg Body weight)



$C_{19}H_{19}N_7O_6$
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



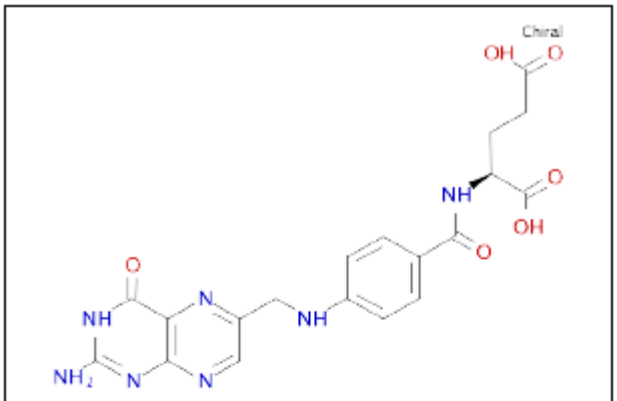
$C_{19}H_{19}N_7O_6$
Molecular Weight: 441.39745
ALogP: -0.232
Rotatable Bonds: 9
Acceptors: 11
Donors: 6

Model Prediction

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

Prediction: Positive if the Bayesian score is above the estimated best cutoff value.
Probability: The esimated 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

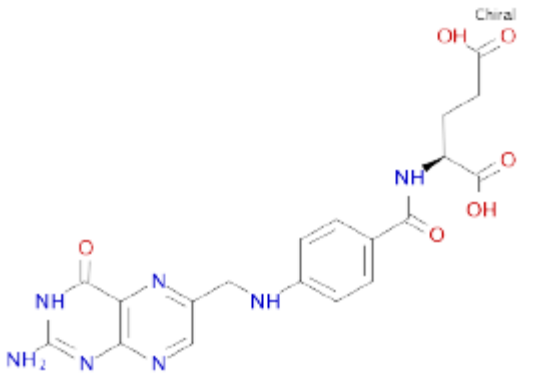


$C_{19}H_{19}N_7O_6$
Molecular Weight: 441.39745
ALogP: -0.232
Rotatable Bonds: 9
Acceptors: 11
Donors: 6

Model Prediction

Prediction: Weak-Sensitizer
Probability: 0.802
Enrichment: 1.035
Bayesian Score: -2.527

Prediction: Positive if the Bayesian score is above the estimated best cutoff value.
Probability: The esimated 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.



$C_{19}H_{19}N_7O_6$
Molecular Weight: 441.39745
ALogP: -0.232
Rotatable Bonds: 9
Acceptors: 11
Donors: 6

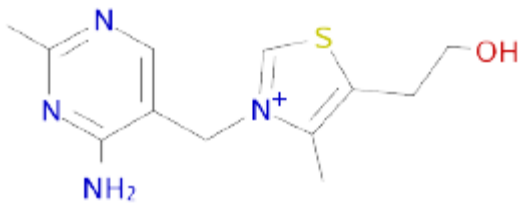
Model Prediction

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 esimated 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



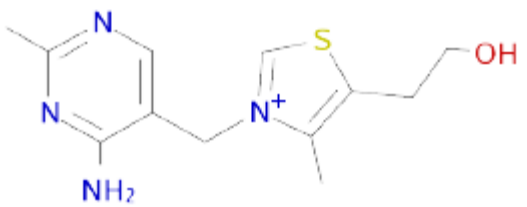
$C_{12}H_{17}N_4OS$
Molecular Weight: 265.35458
ALogP: 1.048
Rotatable Bonds: 4
Acceptors: 4
Donors: 2

Model Prediction

Prediction: Carcinogen
Probability: 0.239
Enrichment: 0.747
Bayesian Score: -0.319

Prediction: Positive if the Bayesian score is above the estimated best cutoff value.
Probability: The esimated 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



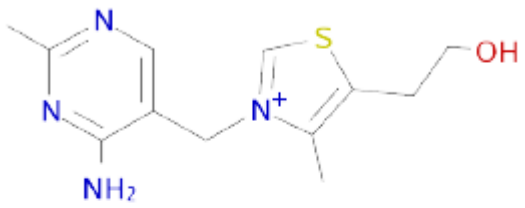
$C_{12}H_{17}N_4OS$
Molecular Weight: 265.35458
ALogP: 1.048
Rotatable Bonds: 4
Acceptors: 4
Donors: 2

Model Prediction

Prediction: Non-Mutagen
Probability: 0.274
Enrichment: 0.490
Bayesian Score: -12.406

Prediction: Positive if the Bayesian score is above the estimated best cutoff value.
Probability: The esimated 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.

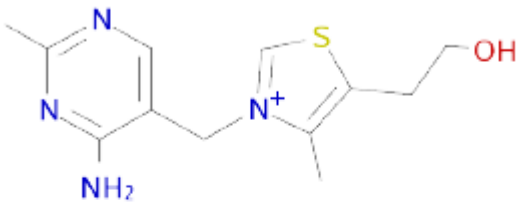
Rat oral LD50 (g/Kg Body weight)



$C_{12}H_{17}N_4OS$
Molecular Weight: 265.35458
ALogP: 1.048
Rotatable Bonds: 4
Acceptors: 4
Donors: 2

Model Prediction
Prediction: 1.308
Unit: g/kg_body_weight

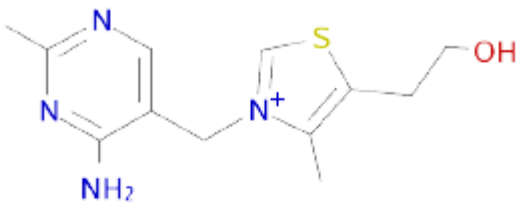
Rat Maximum tolerated dose (g/Kg Body weight)



$C_{12}H_{17}N_4OS$
Molecular Weight: 265.35458
ALogP: 1.048
Rotatable Bonds: 4
Acceptors: 4
Donors: 2

Model Prediction
Prediction: 0.097
Unit: g/kg_body_weight

Skin Irritancy



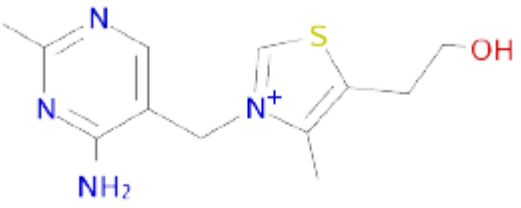
$C_{12}H_{17}N_4OS$
Molecular Weight: 265.35458
ALogP: 1.048
Rotatable Bonds: 4
Acceptors: 4
Donors: 2

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 esimated 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

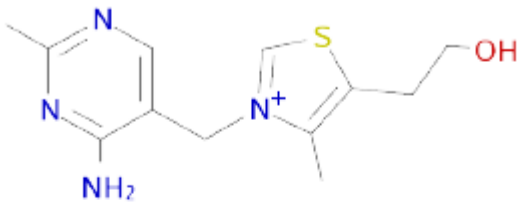


$C_{12}H_{17}N_4OS$
Molecular Weight: 265.35458
ALogP: 1.048
Rotatable Bonds: 4
Acceptors: 4
Donors: 2

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 esimated 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.



$C_{12}H_{17}N_4OS$
Molecular Weight: 265.35458
ALogP: 1.048
Rotatable Bonds: 4
Acceptors: 4
Donors: 2

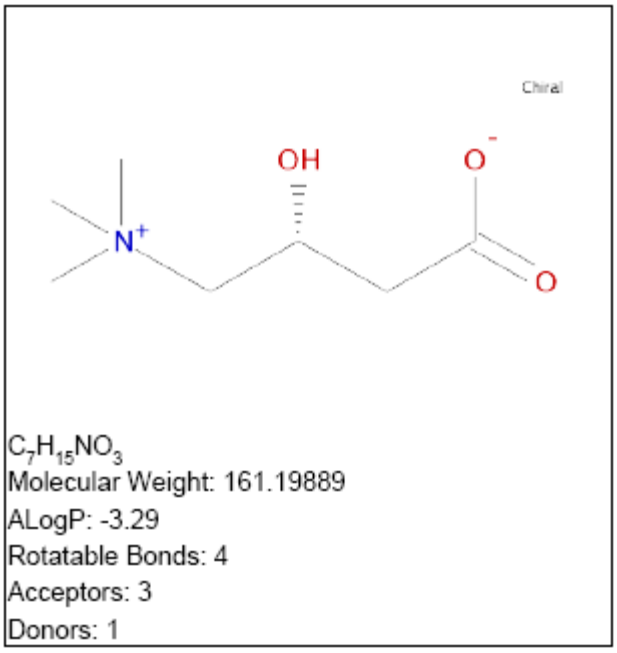
Model Prediction

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 esimated 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)

FDA Rodent Carcinogenicity



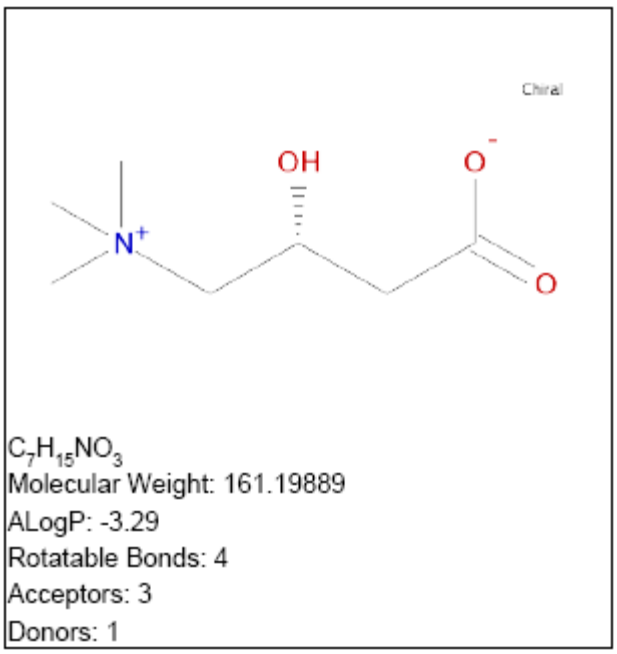
$C_7H_{15}NO_3$
Molecular Weight: 161.19889
ALogP: -3.29
Rotatable Bonds: 4
Acceptors: 3
Donors: 1

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 esimated 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



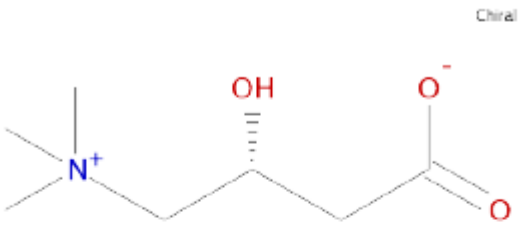
$C_7H_{15}NO_3$
Molecular Weight: 161.19889
ALogP: -3.29
Rotatable Bonds: 4
Acceptors: 3
Donors: 1

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 esimated 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.

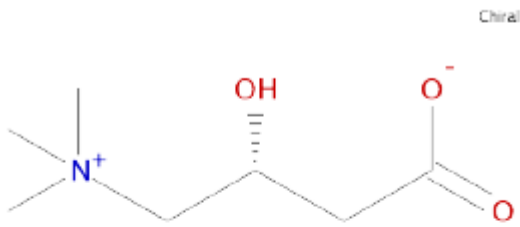
Rat oral LD50 (g/Kg Body weight)



$C_7H_{15}NO_3$
Molecular Weight: 161.19889
ALogP: -3.29
Rotatable Bonds: 4
Acceptors: 3
Donors: 1

Model Prediction
Prediction: 1.101
Unit: g/kg_body_weight

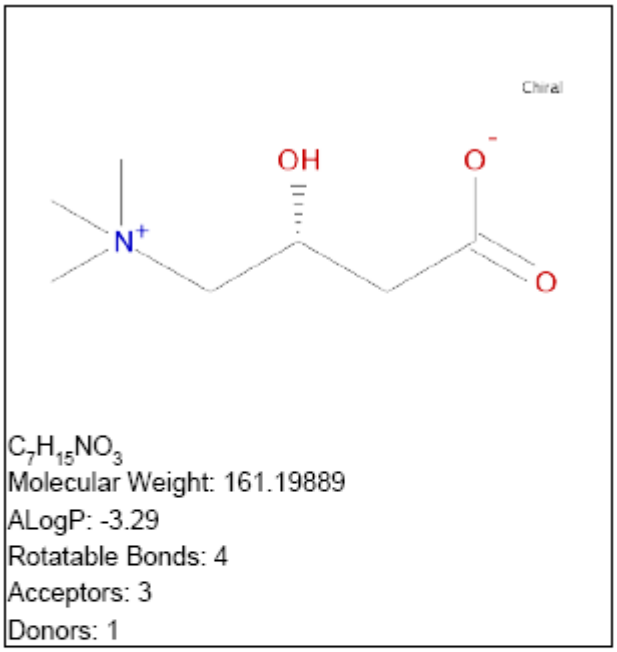
Rat Maximum tolerated dose (g/Kg Body weight)



$C_7H_{15}NO_3$
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



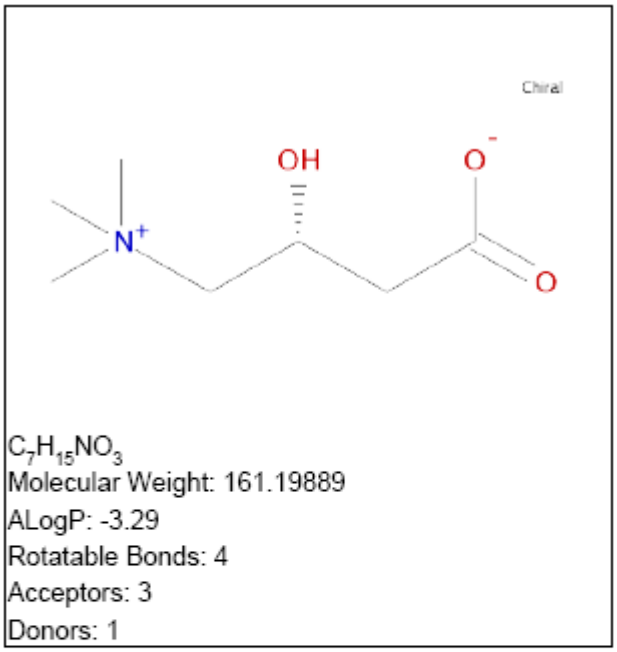
$C_7H_{15}NO_3$
Molecular Weight: 161.19889
ALogP: -3.29
Rotatable Bonds: 4
Acceptors: 3
Donors: 1

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 esimated 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



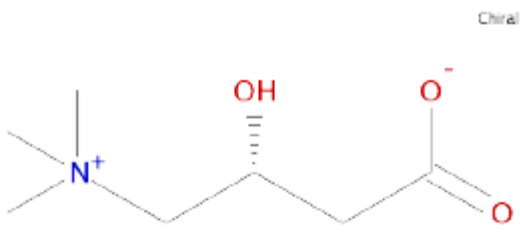
$C_7H_{15}NO_3$
Molecular Weight: 161.19889
ALogP: -3.29
Rotatable Bonds: 4
Acceptors: 3
Donors: 1

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 esimated 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



$C_7H_{15}NO_3$
Molecular Weight: 161.19889
ALogP: -3.29
Rotatable Bonds: 4
Acceptors: 3
Donors: 1

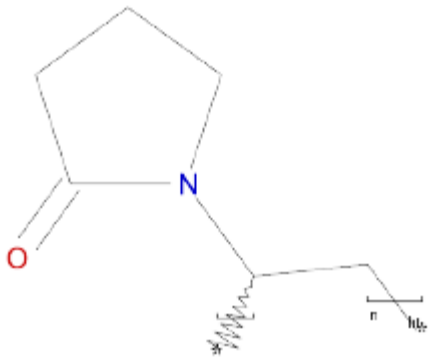
Model Prediction

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 esimated 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



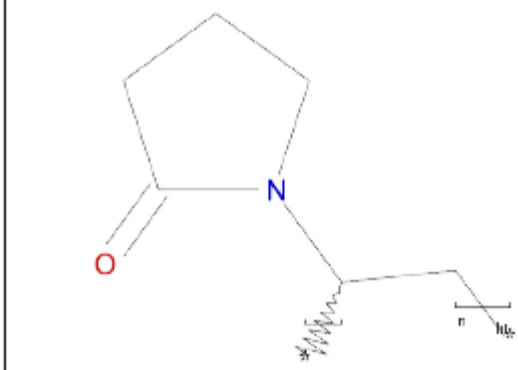
$C_6H_9{}^*{}_2NO$
Molecular Weight: 111.14176
ALogP: 6.1e-002
Rotatable Bonds: 2
Acceptors: 1
Donors: 0

Model Prediction

Prediction: Carcinogen
Probability: 0.256
Enrichment: 0.798
Bayesian Score: 0.693

Prediction: Positive if the Bayesian score is above the estimated best cutoff value.
Probability: The esimated 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



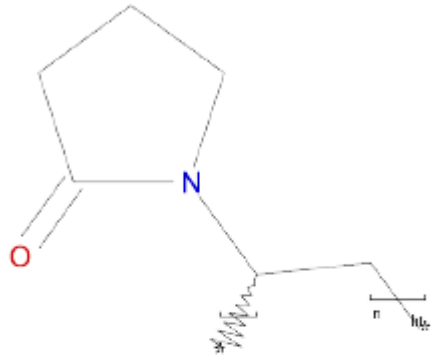
$C_6H_9{}^*{}_2NO$
Molecular Weight: 111.14176
ALogP: 6.1e-002
Rotatable Bonds: 2
Acceptors: 1
Donors: 0

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 esimated 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.

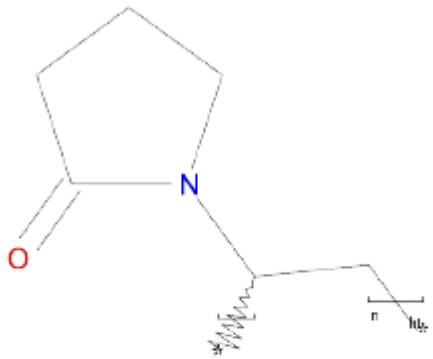
Rat oral LD50 (g/Kg Body weight)



$C_6H_9{}^*{}_2NO$
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

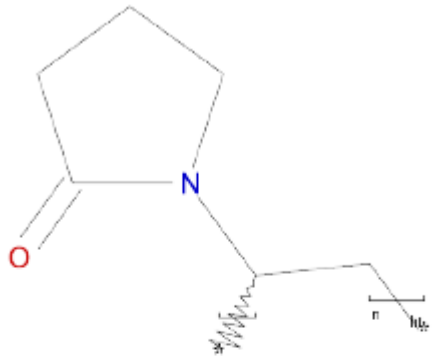
Rat Maximum tolerated dose (g/Kg Body weight)



$C_6H_9{}^*{}_2NO$
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



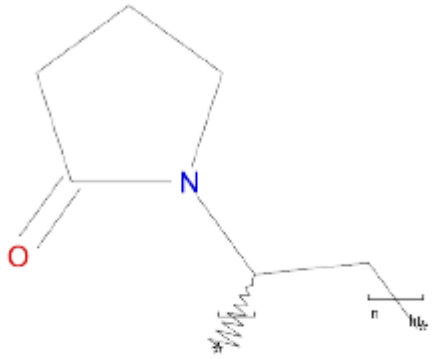
$C_6H_9{}^*{}_2NO$
Molecular Weight: 111.14176
ALogP: 6.1e-002
Rotatable Bonds: 2
Acceptors: 1
Donors: 0

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 esimated 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



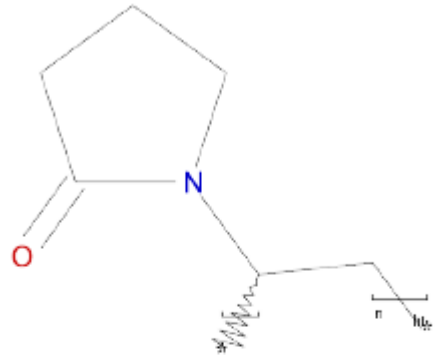
$C_6H_9{}^*{}_2NO$
Molecular Weight: 111.14176
ALogP: 6.1e-002
Rotatable Bonds: 2
Acceptors: 1
Donors: 0

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 esimated 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



$C_6H_9{}^*{}_2NO$
Molecular Weight: 111.14176
ALogP: 6.1e-002
Rotatable Bonds: 2
Acceptors: 1
Donors: 0

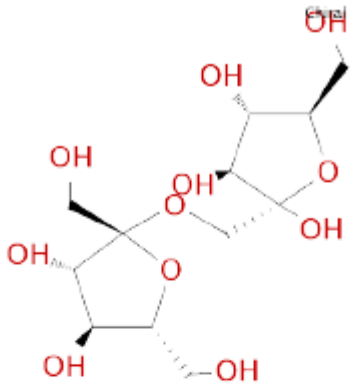
Model Prediction

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 esimated 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



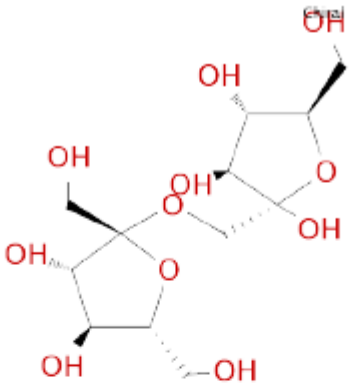
$C_{12}H_{22}O_{11}$
Molecular Weight: 342.29648
ALogP: -4.361
Rotatable Bonds: 6
Acceptors: 11
Donors: 8

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 esimated 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



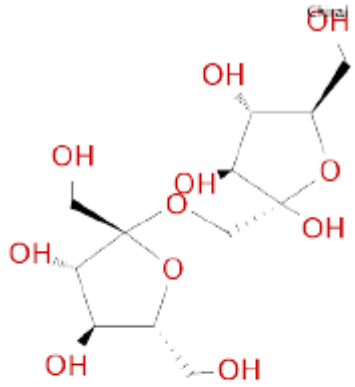
$C_{12}H_{22}O_{11}$
Molecular Weight: 342.29648
ALogP: -4.361
Rotatable Bonds: 6
Acceptors: 11
Donors: 8

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 esimated 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.

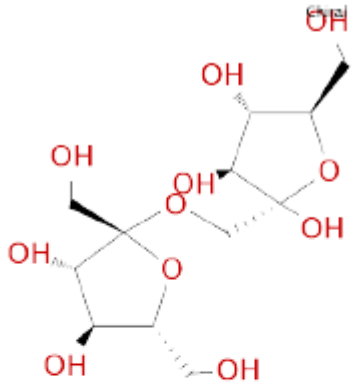
Rat oral LD50 (g/Kg Body weight)



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

Model Prediction
Prediction: 20.789
Unit: g/kg_body_weight

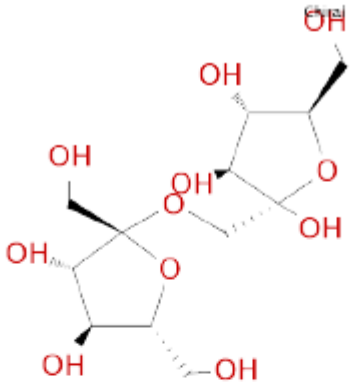
Rat Maximum tolerated dose (g/Kg Body weight)



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

Model Prediction
Prediction: 0.000
Unit: g/kg_body_weight

Skin Irritancy



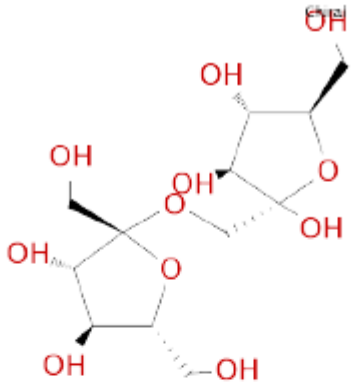
$C_{12}H_{22}O_{11}$
Molecular Weight: 342.29648
ALogP: -4.361
Rotatable Bonds: 6
Acceptors: 11
Donors: 8

Model Prediction

Prediction: Mild
Probability: 0.159
Enrichment: 0.432
Bayesian Score: -6.949

Prediction: Positive if the Bayesian score is above the estimated best cutoff value.
Probability: The esimated 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

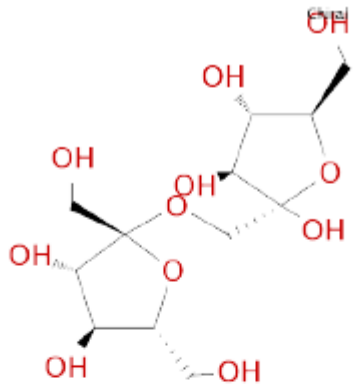


$C_{12}H_{22}O_{11}$
Molecular Weight: 342.29648
ALogP: -4.361
Rotatable Bonds: 6
Acceptors: 11
Donors: 8

Model Prediction

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

Prediction: Positive if the Bayesian score is above the estimated best cutoff value.
Probability: The esimated 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.



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

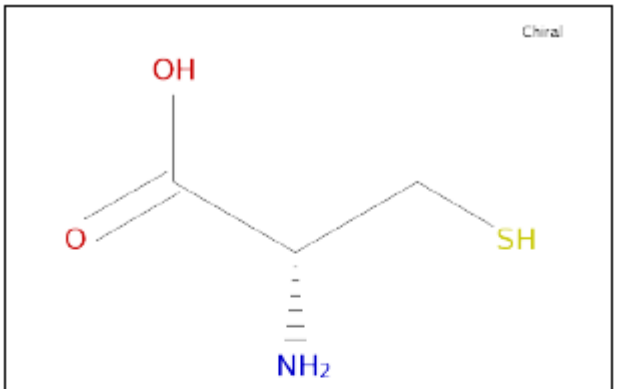
Model Prediction

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 esimated 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



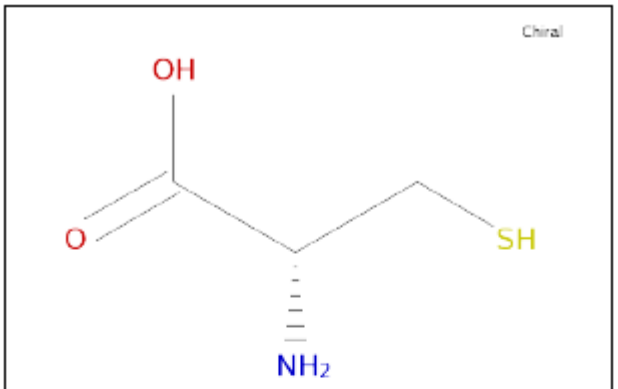
$C_3H_7NO_2S$
Molecular Weight: 121.15818
ALogP: -3.078
Rotatable Bonds: 2
Acceptors: 4
Donors: 3

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 esimated 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



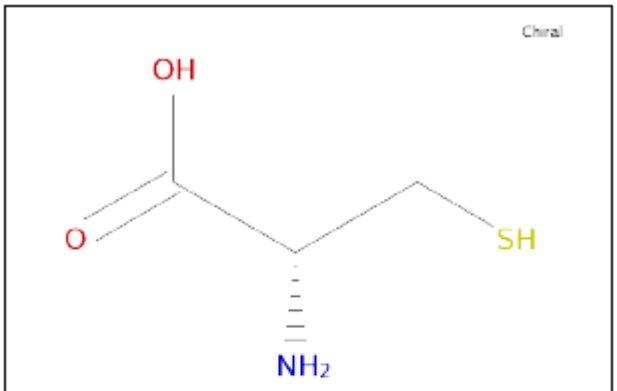
$C_3H_7NO_2S$
Molecular Weight: 121.15818
ALogP: -3.078
Rotatable Bonds: 2
Acceptors: 4
Donors: 3

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 esimated 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.

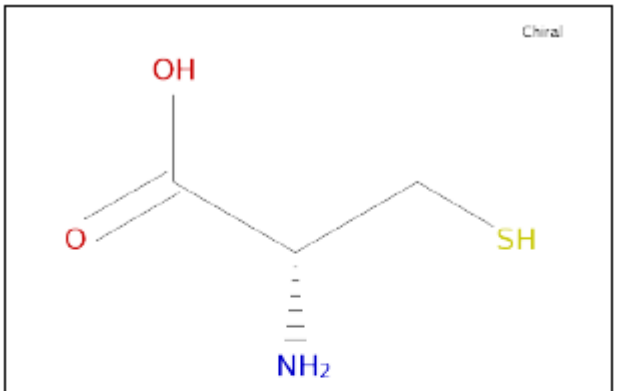
Rat oral LD50 (g/Kg Body weight)



$C_3H_7NO_2S$
Molecular Weight: 121.15818
ALogP: -3.078
Rotatable Bonds: 2
Acceptors: 4
Donors: 3

Model Prediction
Prediction: 0.514
Unit: g/kg_body_weight

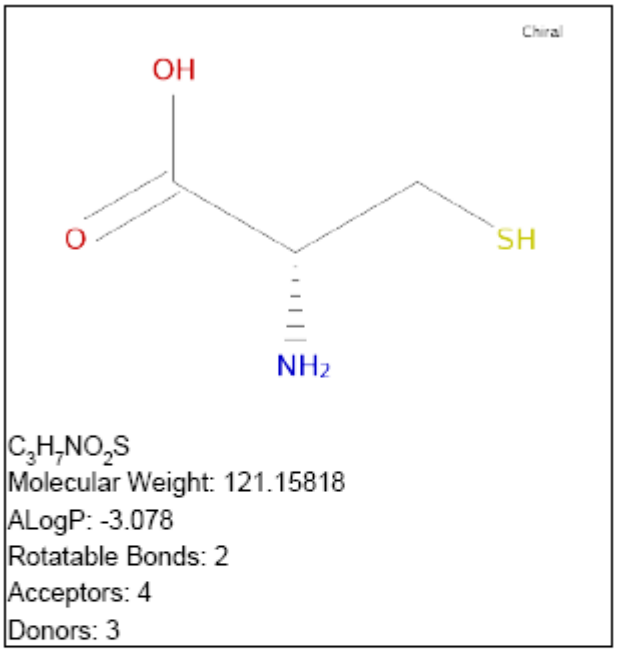
Rat Maximum tolerated dose (g/Kg Body weight)



$C_3H_7NO_2S$
Molecular Weight: 121.15818
ALogP: -3.078
Rotatable Bonds: 2
Acceptors: 4
Donors: 3

Model Prediction
Prediction: 0.653
Unit: g/kg_body_weight

Skin Irritancy



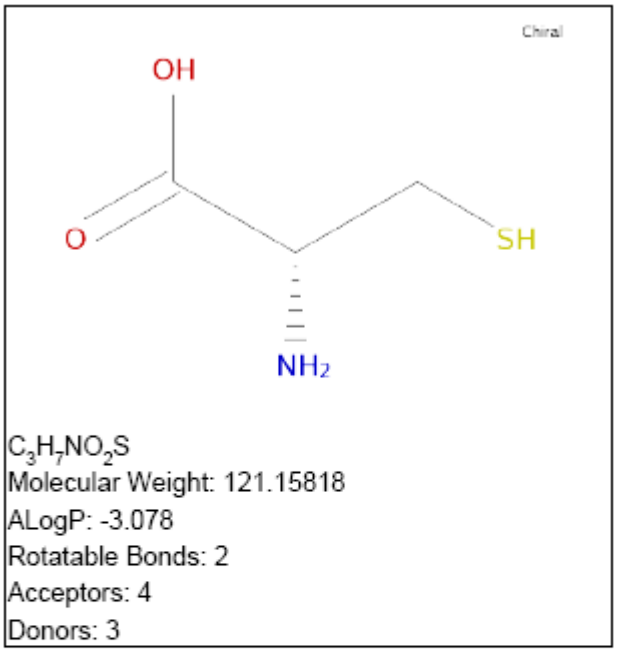
$C_3H_7NO_2S$
Molecular Weight: 121.15818
ALogP: -3.078
Rotatable Bonds: 2
Acceptors: 4
Donors: 3

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 esimated 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



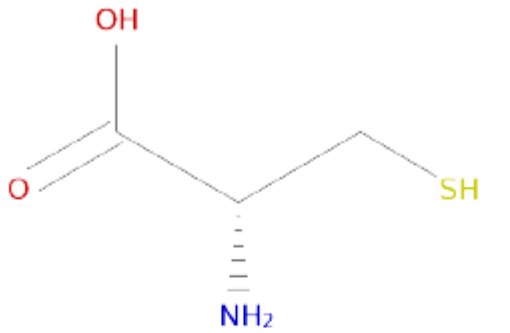
$C_3H_7NO_2S$
Molecular Weight: 121.15818
ALogP: -3.078
Rotatable Bonds: 2
Acceptors: 4
Donors: 3

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 esimated 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



$C_3H_7NO_2S$
Molecular Weight: 121.15818
ALogP: -3.078
Rotatable Bonds: 2
Acceptors: 4
Donors: 3

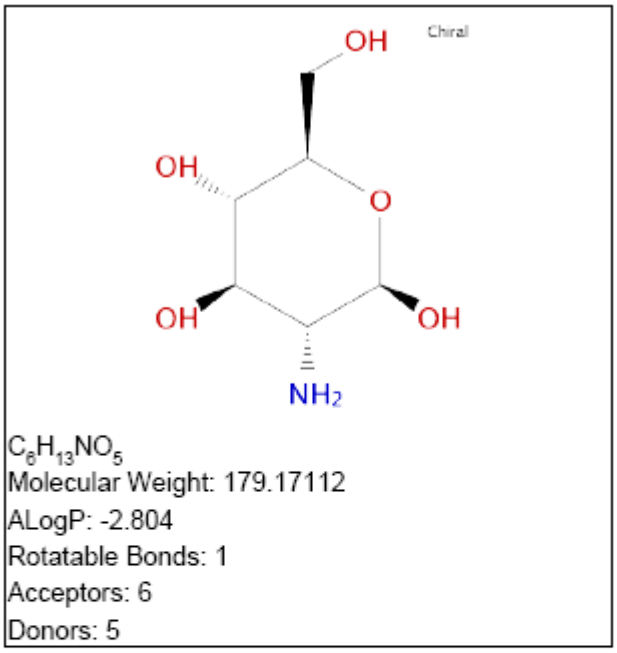
Model Prediction

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 esimated 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



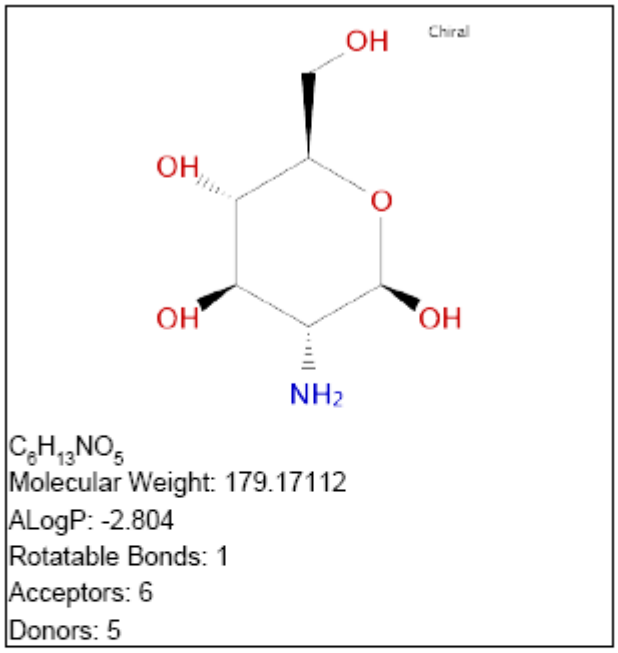
$C_6H_{13}NO_5$
Molecular Weight: 179.17112
ALogP: -2.804
Rotatable Bonds: 1
Acceptors: 6
Donors: 5

Model Prediction

Prediction: Non-Carcinogen
Probability: 0.215
Enrichment: 0.672
Bayesian Score: -2.754

Prediction: Positive if the Bayesian score is above the estimated best cutoff value.
Probability: The esimated 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



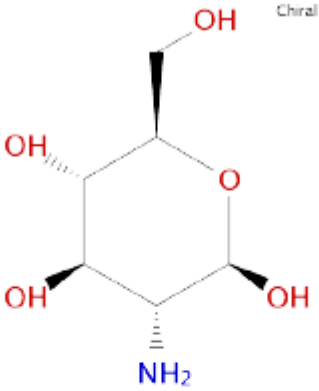
$C_6H_{13}NO_5$
Molecular Weight: 179.17112
ALogP: -2.804
Rotatable Bonds: 1
Acceptors: 6
Donors: 5

Model Prediction

Prediction: Non-Mutagen
Probability: 0.711
Enrichment: 1.273
Bayesian Score: -1.342

Prediction: Positive if the Bayesian score is above the estimated best cutoff value.
Probability: The esimated 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.

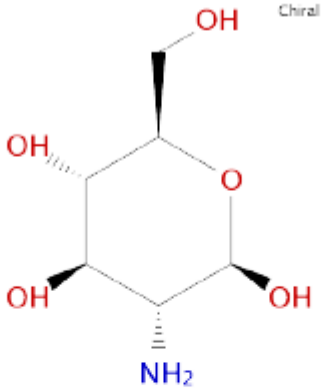
Rat oral LD50 (g/Kg Body weight)



$C_6H_{13}NO_5$
Molecular Weight: 179.17112
ALogP: -2.804
Rotatable Bonds: 1
Acceptors: 6
Donors: 5

Model Prediction
Prediction: 3.241
Unit: g/kg_body_weight

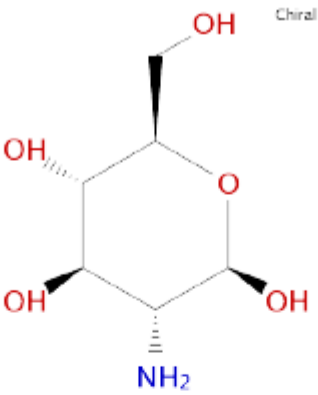
Rat Maximum tolerated dose (g/Kg Body weight)



$C_6H_{13}NO_5$
Molecular Weight: 179.17112
ALogP: -2.804
Rotatable Bonds: 1
Acceptors: 6
Donors: 5

Model Prediction
Prediction: 0.268
Unit: g/kg_body_weight

Skin Irritancy



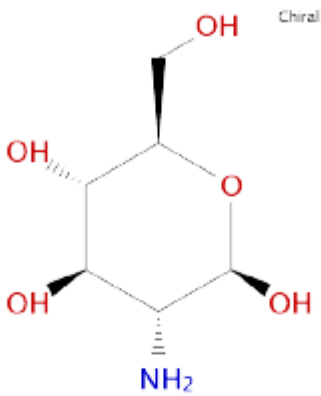
$C_6H_{13}NO_5$
Molecular Weight: 179.17112
ALogP: -2.804
Rotatable Bonds: 1
Acceptors: 6
Donors: 5

Model Prediction

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

Prediction: Positive if the Bayesian score is above the estimated best cutoff value.
Probability: The esimated 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

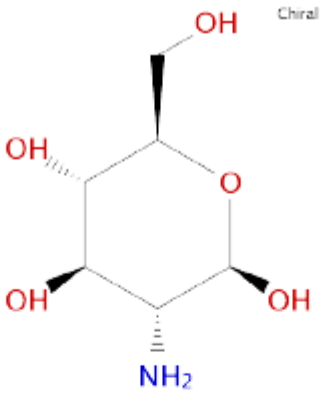


$C_6H_{13}NO_5$
Molecular Weight: 179.17112
ALogP: -2.804
Rotatable Bonds: 1
Acceptors: 6
Donors: 5

Model Prediction

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

Prediction: Positive if the Bayesian score is above the estimated best cutoff value.
Probability: The esimated 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.



$C_6H_{13}NO_5$
Molecular Weight: 179.17112
ALogP: -2.804
Rotatable Bonds: 1
Acceptors: 6
Donors: 5

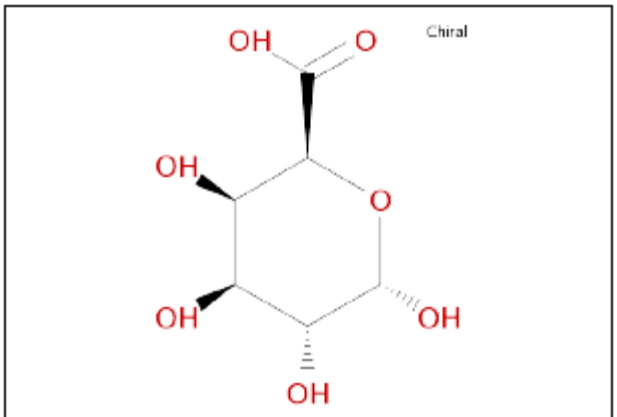
Model Prediction

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 esimated 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



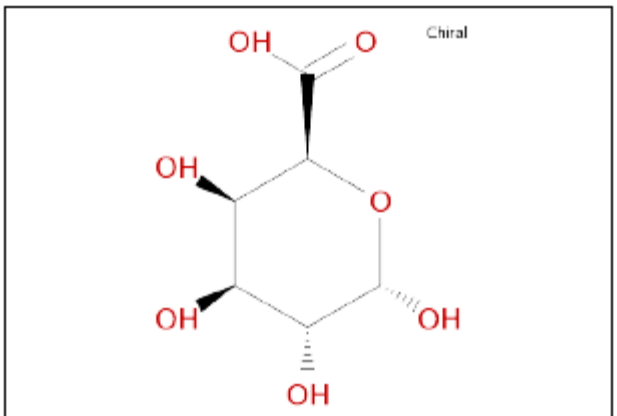
$C_6H_{10}O_7$
Molecular Weight: 194.1394
ALogP: -2.386
Rotatable Bonds: 1
Acceptors: 7
Donors: 5

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 esimated 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



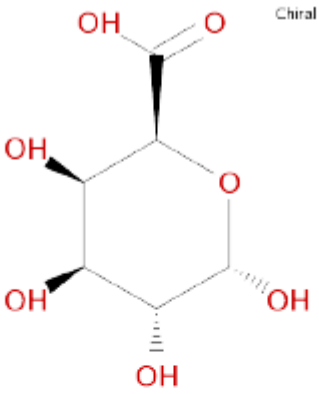
$C_6H_{10}O_7$
Molecular Weight: 194.1394
ALogP: -2.386
Rotatable Bonds: 1
Acceptors: 7
Donors: 5

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 esimated 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.

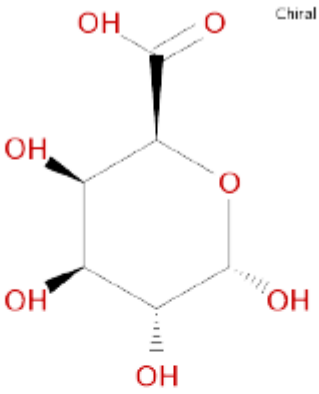
Rat oral LD50 (g/Kg Body weight)



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

Model Prediction
Prediction: 0.525
Unit: g/kg_body_weight

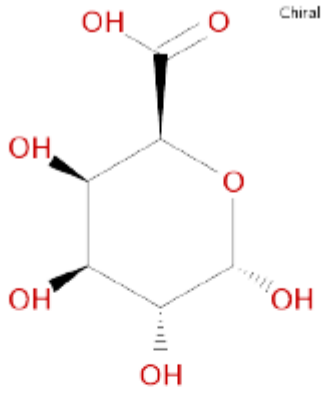
Rat Maximum tolerated dose (g/Kg Body weight)



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

Model Prediction
Prediction: 3.576
Unit: g/kg_body_weight

Skin Irritancy



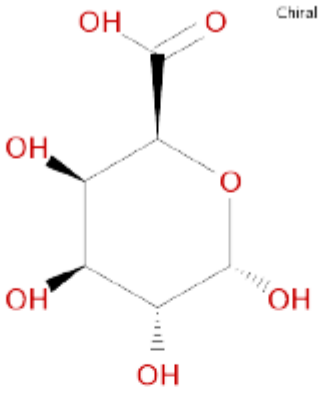
$C_6H_{10}O_7$
Molecular Weight: 194.1394
ALogP: -2.386
Rotatable Bonds: 1
Acceptors: 7
Donors: 5

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 esimated 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



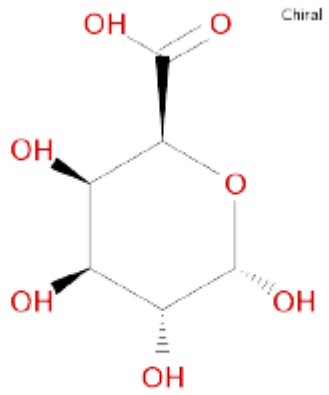
$C_6H_{10}O_7$
Molecular Weight: 194.1394
ALogP: -2.386
Rotatable Bonds: 1
Acceptors: 7
Donors: 5

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 esimated 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



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

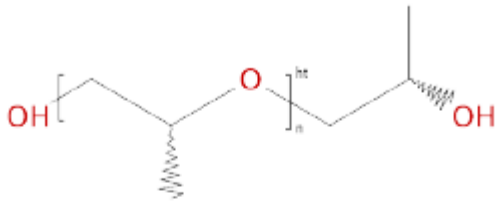
Model Prediction

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 esimated 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



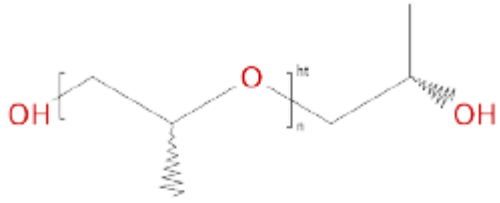
$C_6H_{14}O_3$
Molecular Weight: 134.17356
ALogP: -0.274
Rotatable Bonds: 4
Acceptors: 3
Donors: 2

Model Prediction

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

Prediction: Positive if the Bayesian score is above the estimated best cutoff value.
Probability: The esimated 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



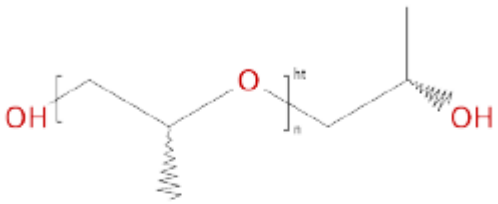
$C_6H_{14}O_3$
Molecular Weight: 134.17356
ALogP: -0.274
Rotatable Bonds: 4
Acceptors: 3
Donors: 2

Model Prediction

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

Prediction: Positive if the Bayesian score is above the estimated best cutoff value.
Probability: The esimated 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.

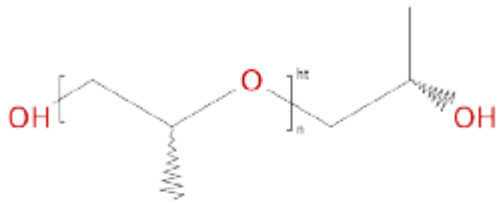
Rat oral LD50 (g/Kg Body weight)



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

Model Prediction
Prediction: 12.098
Unit: g/kg_body_weight

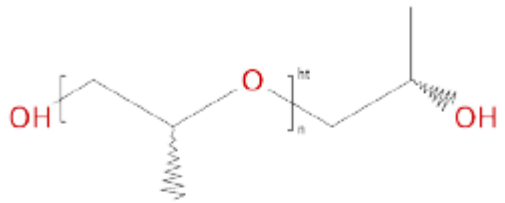
Rat Maximum tolerated dose (g/Kg Body weight)



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

Model Prediction
Prediction: 0.187
Unit: g/kg_body_weight

Skin Irritancy



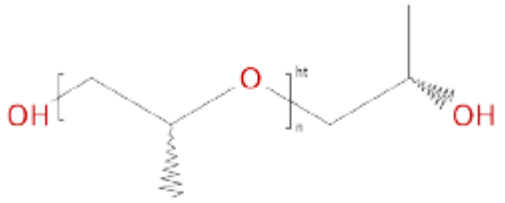
$C_6H_{14}O_3$
Molecular Weight: 134.17356
ALogP: -0.274
Rotatable Bonds: 4
Acceptors: 3
Donors: 2

Model Prediction

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

Prediction: Positive if the Bayesian score is above the estimated best cutoff value.
Probability: The esimated 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



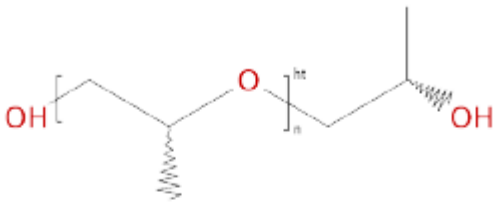
$C_6H_{14}O_3$
Molecular Weight: 134.17356
ALogP: -0.274
Rotatable Bonds: 4
Acceptors: 3
Donors: 2

Model Prediction

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 esimated 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



$C_6H_{14}O_3$
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 esimated 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

$C_3H_7{}^*{}_2N$
Molecular Weight: 57.09438
ALogP: 0.131
Rotatable Bonds: 3
Acceptors: 1
Donors: 1

Model Prediction

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

Prediction: Positive if the Bayesian score is above the estimated best cutoff value.
Probability: The esimated 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

$C_3H_7{}^*{}_2N$
Molecular Weight: 57.09438
ALogP: 0.131
Rotatable Bonds: 3
Acceptors: 1
Donors: 1

Model Prediction

Prediction: Non-Mutagen
Probability: 0.666
Enrichment: 1.192
Bayesian Score: -3.064

Prediction: Positive if the Bayesian score is above the estimated best cutoff value.
Probability: The esimated 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)

$C_3H_7{}^*{}_2N$
Molecular Weight: 57.09438
ALogP: 0.131
Rotatable Bonds: 3
Acceptors: 1
Donors: 1

Model Prediction
Prediction: 0.055
Unit: g/kg_body_weight

Rat Maximum tolerated dose (g/Kg Body weight)

$C_3H_7{}^*{}_2N$
Molecular Weight: 57.09438
ALogP: 0.131
Rotatable Bonds: 3
Acceptors: 1
Donors: 1

Model Prediction
Prediction: 0.089
Unit: g/kg_body_weight

Skin Irritancy

$C_3H_7{}^*{}_2N$
Molecular Weight: 57.09438
ALogP: 0.131
Rotatable Bonds: 3
Acceptors: 1
Donors: 1

Model Prediction

Prediction: Mild
Probability: 0.275
Enrichment: 0.748
Bayesian Score: -3.519

Prediction: Positive if the Bayesian score is above the estimated best cutoff value.
Probability: The esimated 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

$C_3H_7{}^*{}_2N$
Molecular Weight: 57.09438
ALogP: 0.131
Rotatable Bonds: 3
Acceptors: 1
Donors: 1

Model Prediction

Prediction: Non-Sensitizer
Probability: 0.681
Enrichment: 0.993
Bayesian Score: -2.202

Prediction: Positive if the Bayesian score is above the estimated best cutoff value.
Probability: The esimated 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

$C_3H_7{}^*{}_2N$
Molecular Weight: 57.09438
ALogP: 0.131
Rotatable Bonds: 3
Acceptors: 1
Donors: 1

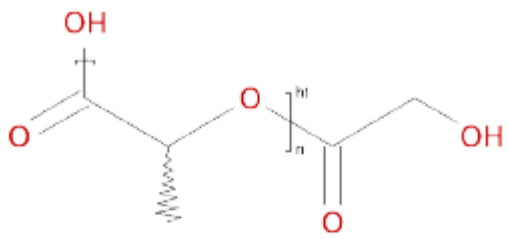
Model Prediction

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 esimated 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



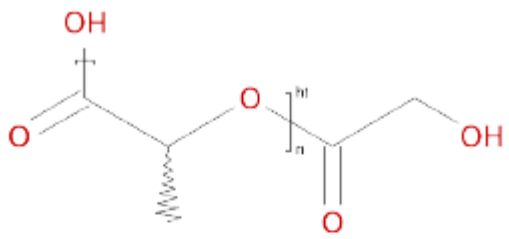
$C_5H_8O_5$
Molecular Weight: 148.11402
ALogP: -0.686
Rotatable Bonds: 4
Acceptors: 5
Donors: 2

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 esimated 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



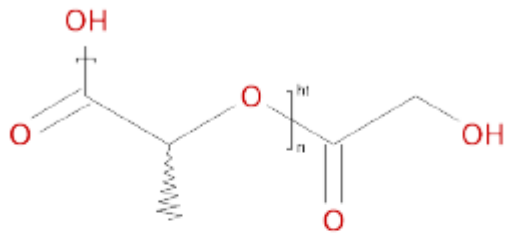
$C_5H_8O_5$
Molecular Weight: 148.11402
ALogP: -0.686
Rotatable Bonds: 4
Acceptors: 5
Donors: 2

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 esimated 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.

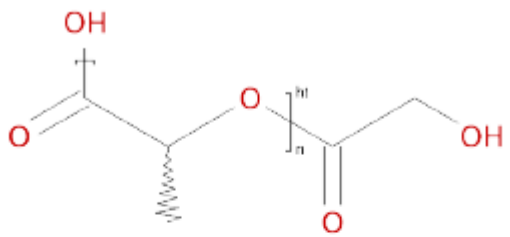
Rat oral LD50 (g/Kg Body weight)



$C_5H_8O_5$
Molecular Weight: 148.11402
ALogP: -0.686
Rotatable Bonds: 4
Acceptors: 5
Donors: 2

Model Prediction
Prediction: 2.982
Unit: g/kg_body_weight

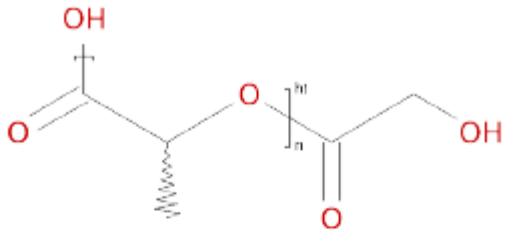
Rat Maximum tolerated dose (g/Kg Body weight)



$C_5H_8O_5$
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



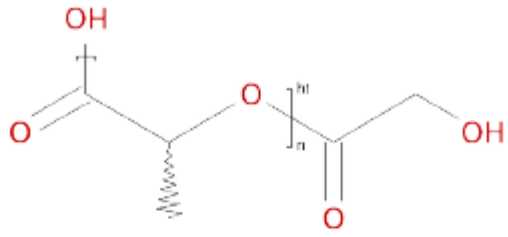
$C_5H_8O_5$
Molecular Weight: 148.11402
ALogP: -0.686
Rotatable Bonds: 4
Acceptors: 5
Donors: 2

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 esimated 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



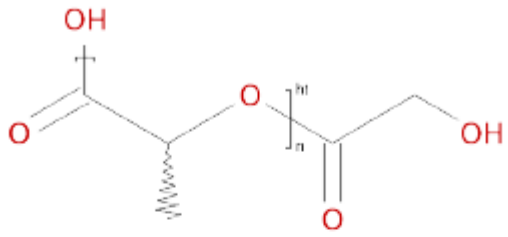
$C_5H_8O_5$
Molecular Weight: 148.11402
ALogP: -0.686
Rotatable Bonds: 4
Acceptors: 5
Donors: 2

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 esimated 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



$C_5H_8O_5$
Molecular Weight: 148.11402
ALogP: -0.686
Rotatable Bonds: 4
Acceptors: 5
Donors: 2

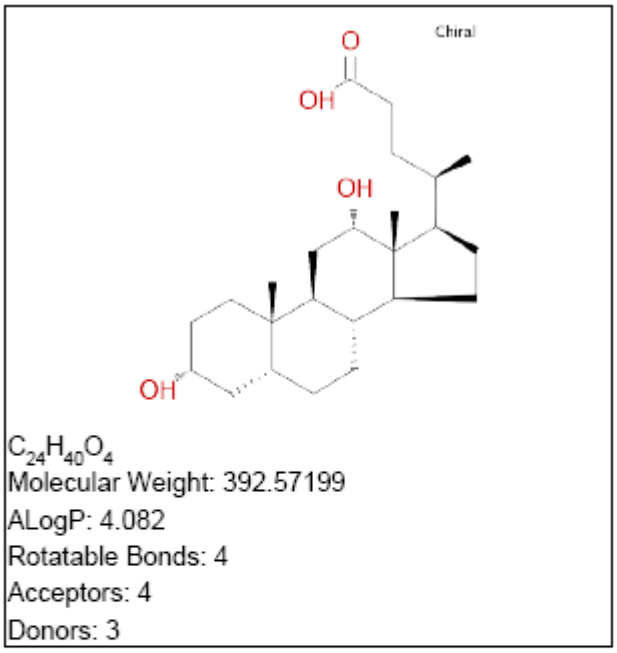
Model Prediction

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 esimated 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



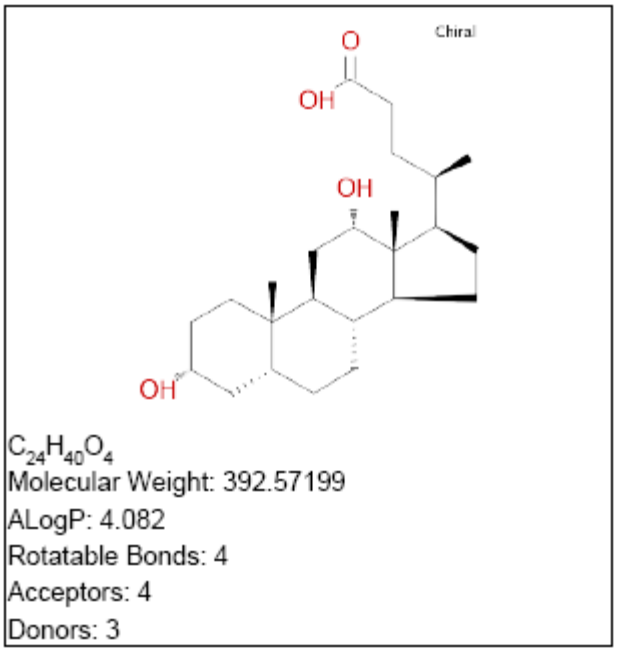
$C_{24}H_{40}O_4$
Molecular Weight: 392.57199
ALogP: 4.082
Rotatable Bonds: 4
Acceptors: 4
Donors: 3

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 esimated 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



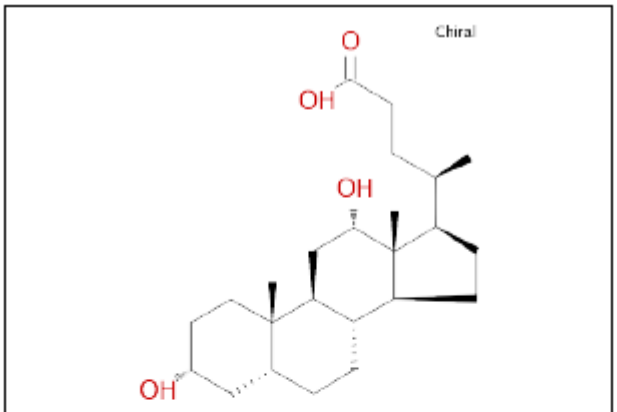
$C_{24}H_{40}O_4$
Molecular Weight: 392.57199
ALogP: 4.082
Rotatable Bonds: 4
Acceptors: 4
Donors: 3

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 esimated 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.

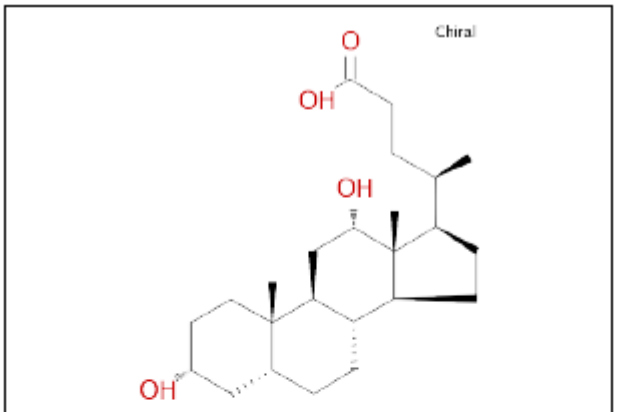
Rat oral LD50 (g/Kg Body weight)



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

Model Prediction
Prediction: 6.358
Unit: g/kg_body_weight

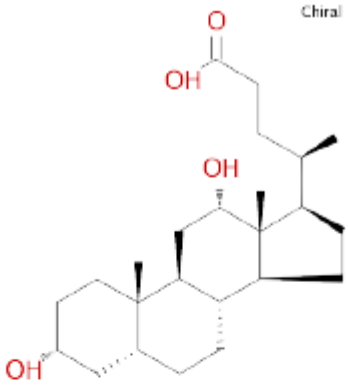
Rat Maximum tolerated dose (g/Kg Body weight)



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

Model Prediction
Prediction: 0.190
Unit: g/kg_body_weight

Skin Irritancy



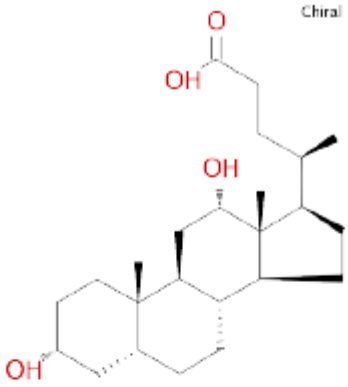
$C_{24}H_{40}O_4$
Molecular Weight: 392.57199
ALogP: 4.082
Rotatable Bonds: 4
Acceptors: 4
Donors: 3

Model Prediction

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 esimated 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



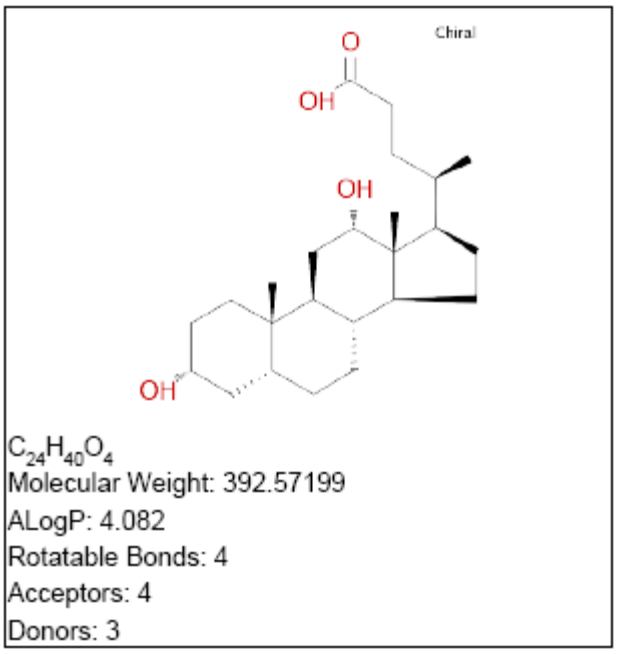
$C_{24}H_{40}O_4$
Molecular Weight: 392.57199
ALogP: 4.082
Rotatable Bonds: 4
Acceptors: 4
Donors: 3

Model Prediction

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 esimated 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

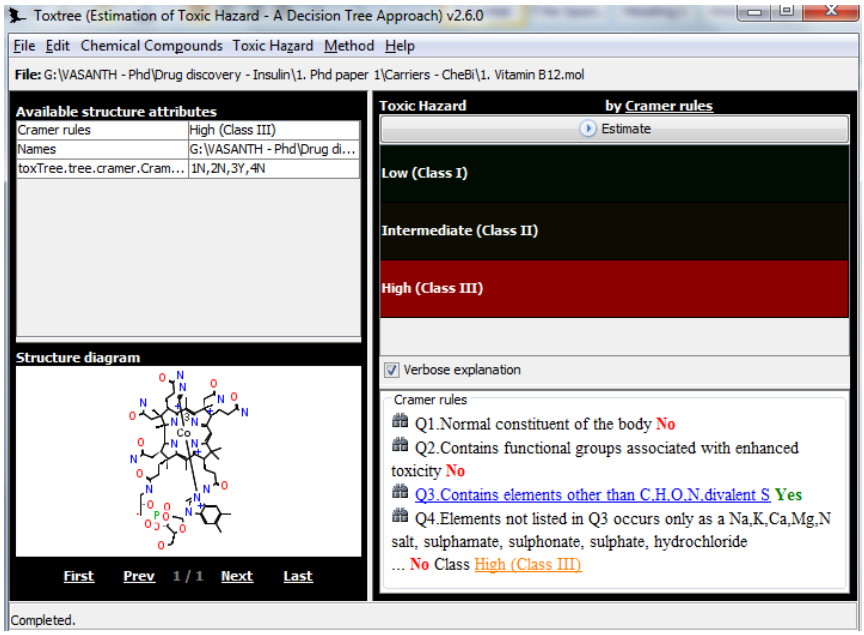
Prediction: Degradable
Probability: 0.837
Enrichment: 1.919
Bayesian Score: 8.551

Prediction: Positive if the Bayesian score is above the estimated best cutoff value.
Probability: The esimated 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.

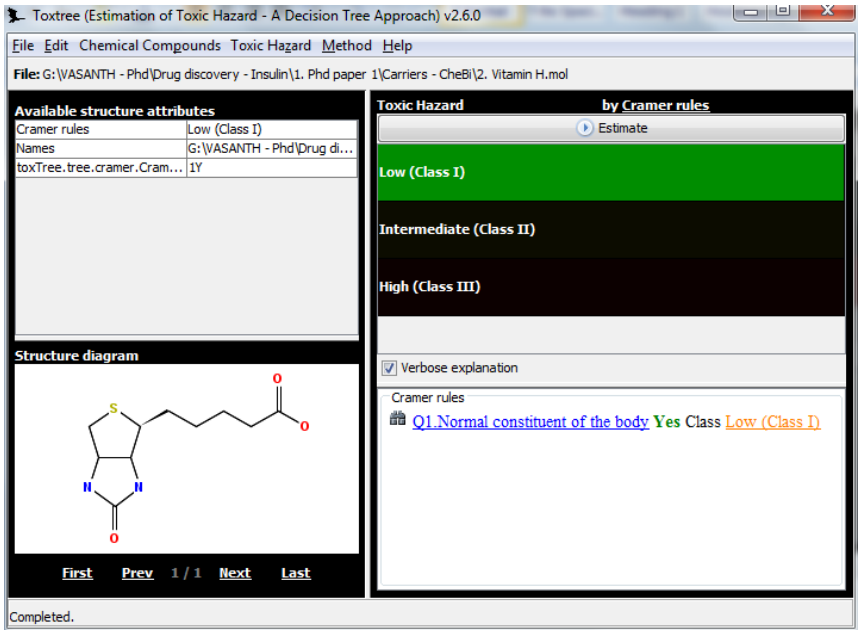
Supplementary figure 4

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 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

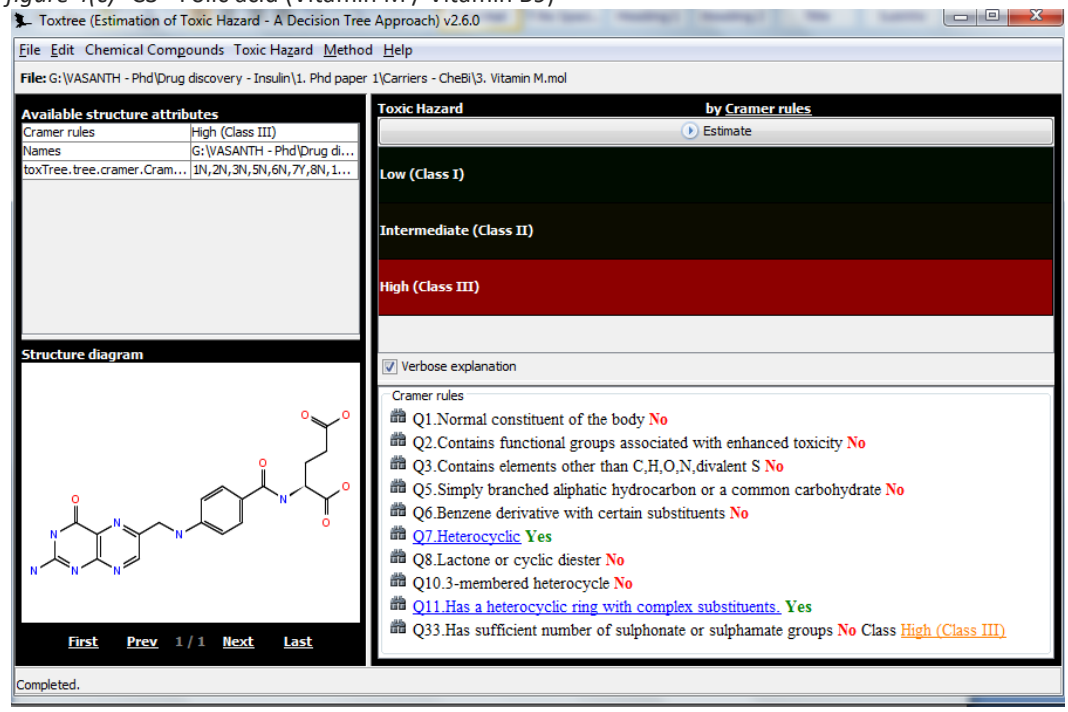
Supplementary figure 4(a)- C1 - Vitamin B12 (cobalamin)



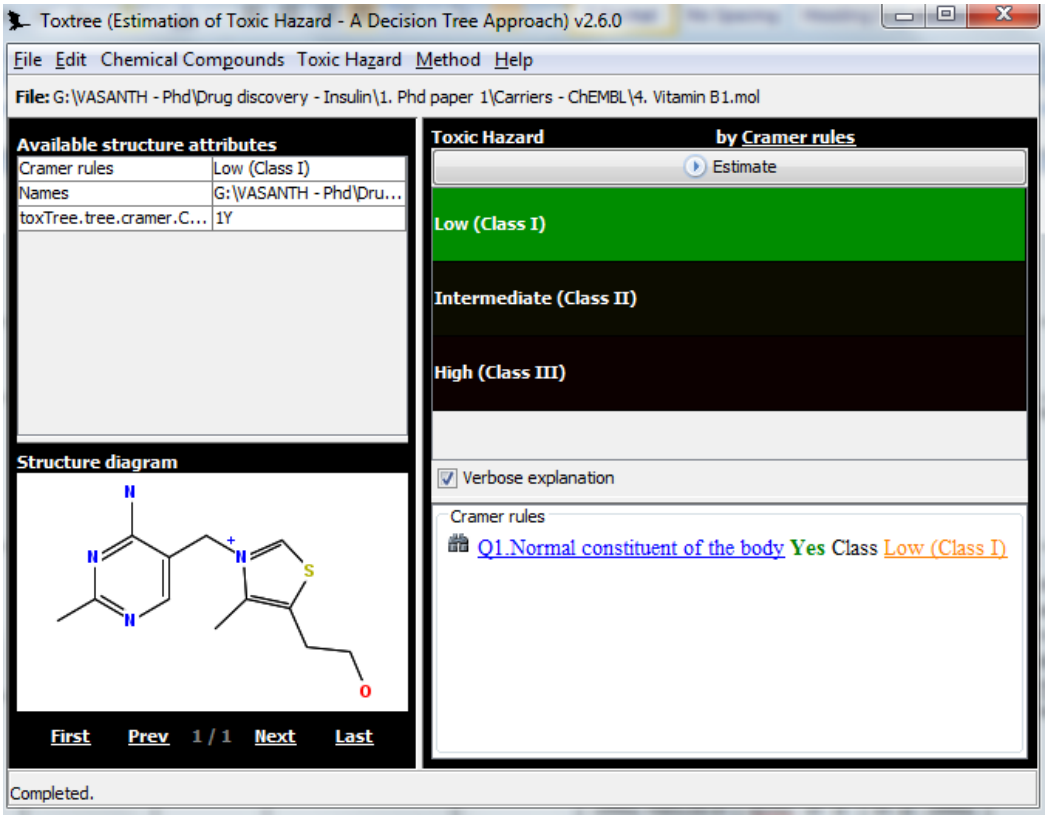
Supplementary figure 4(b)- C2 - Vitamin H (Biotin)



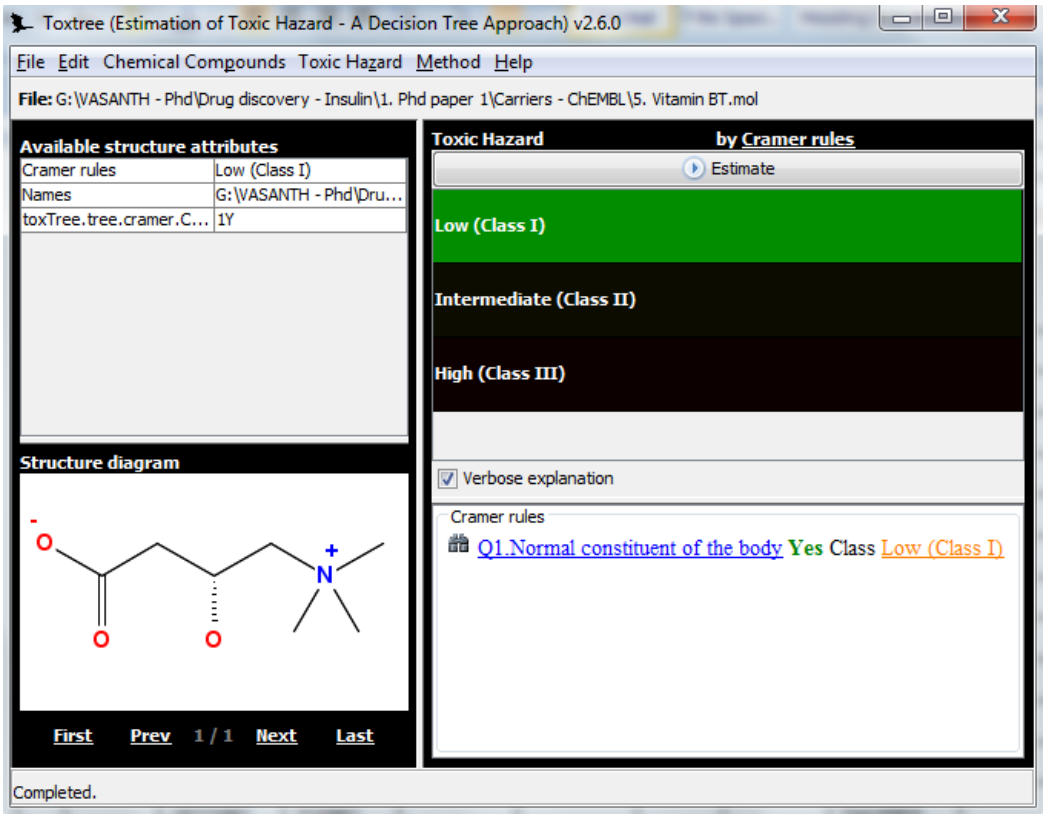
Supplementary figure 4(c)- C3 - Folic acid (Vitamin M / Vitamin B9)



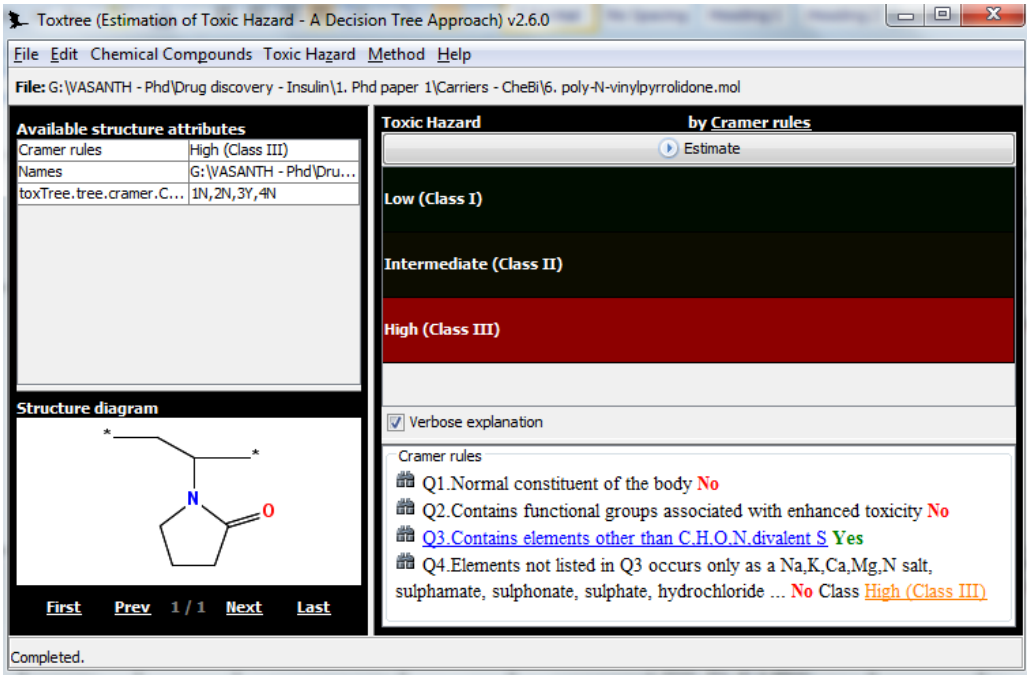
Supplementary figure 4(d)- C4 - Vitamin B1 (Thiamin)



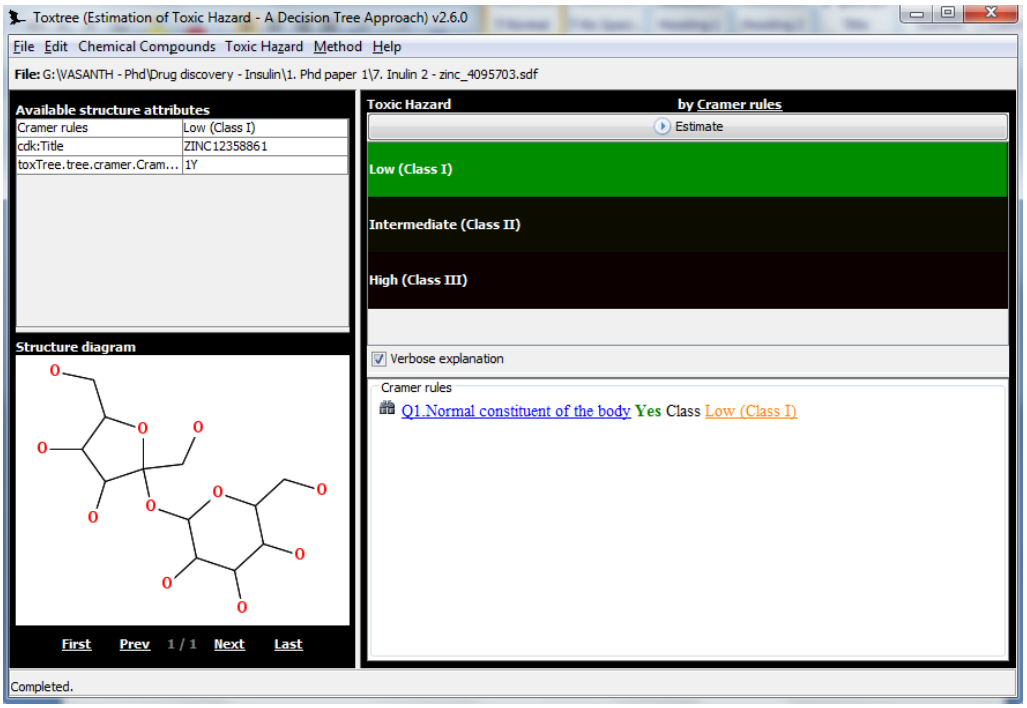
Supplementary figure 4(e)- C5 - L-Carnitine (Vitamin BT)



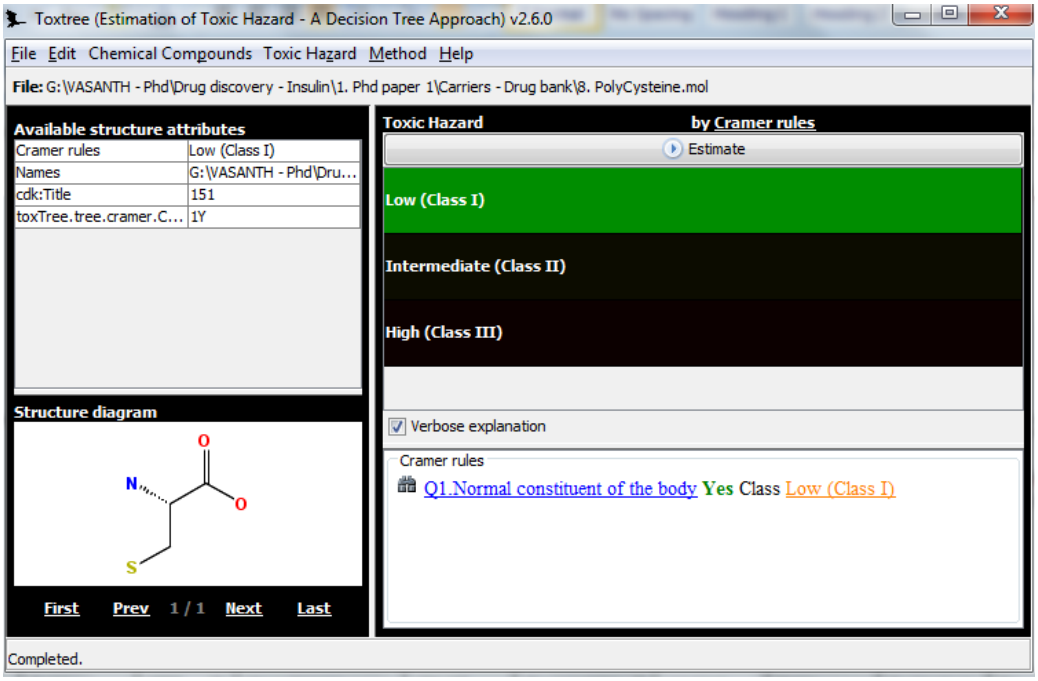
Supplementary figure 4(f)- C6 - Poly-N-vinylpyrrolidone



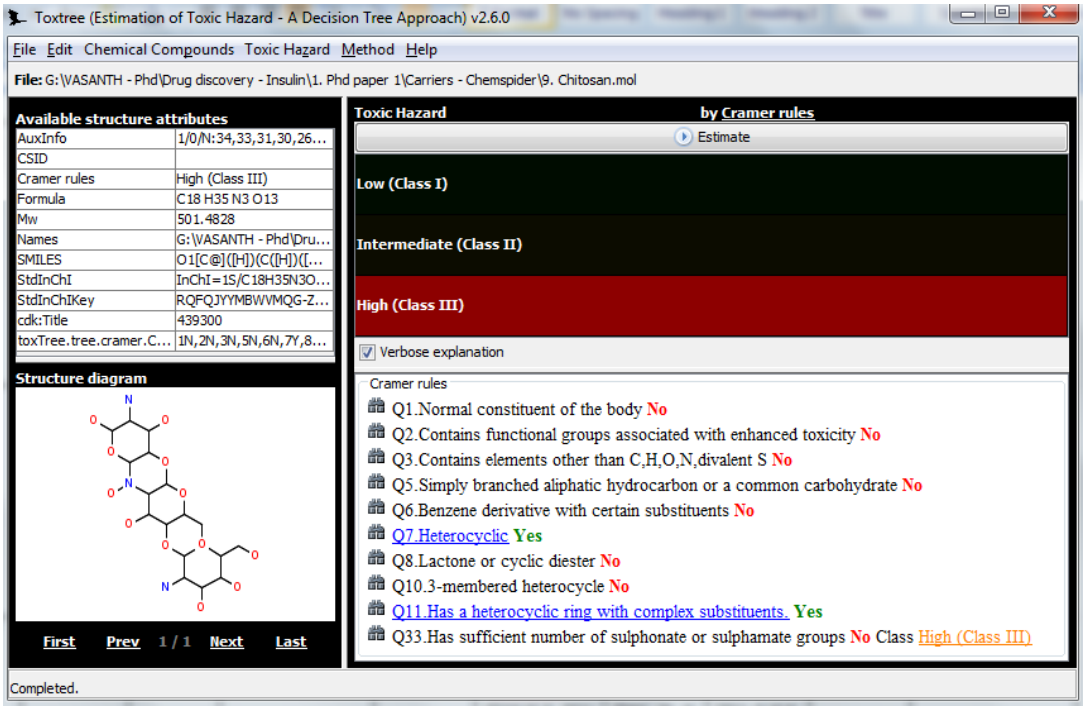
Supplementary figure 4(g)- C7; Inulin



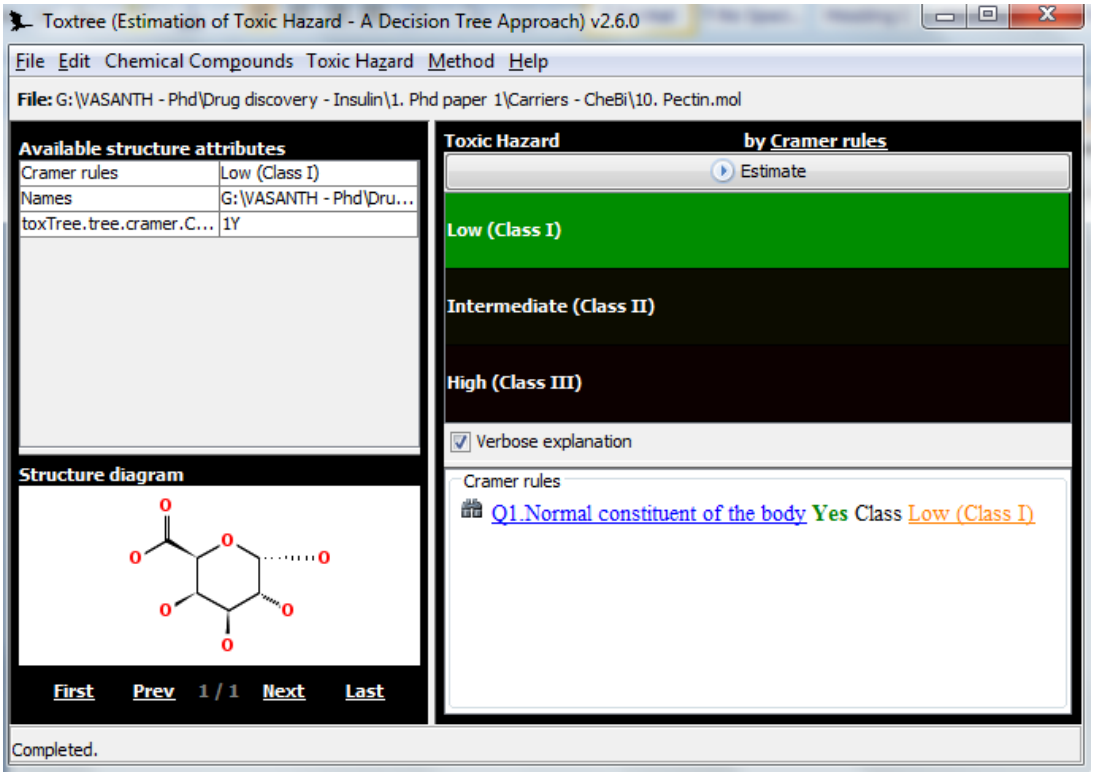
Supplementary figure 3(h)- C8 - Poly Cysteine



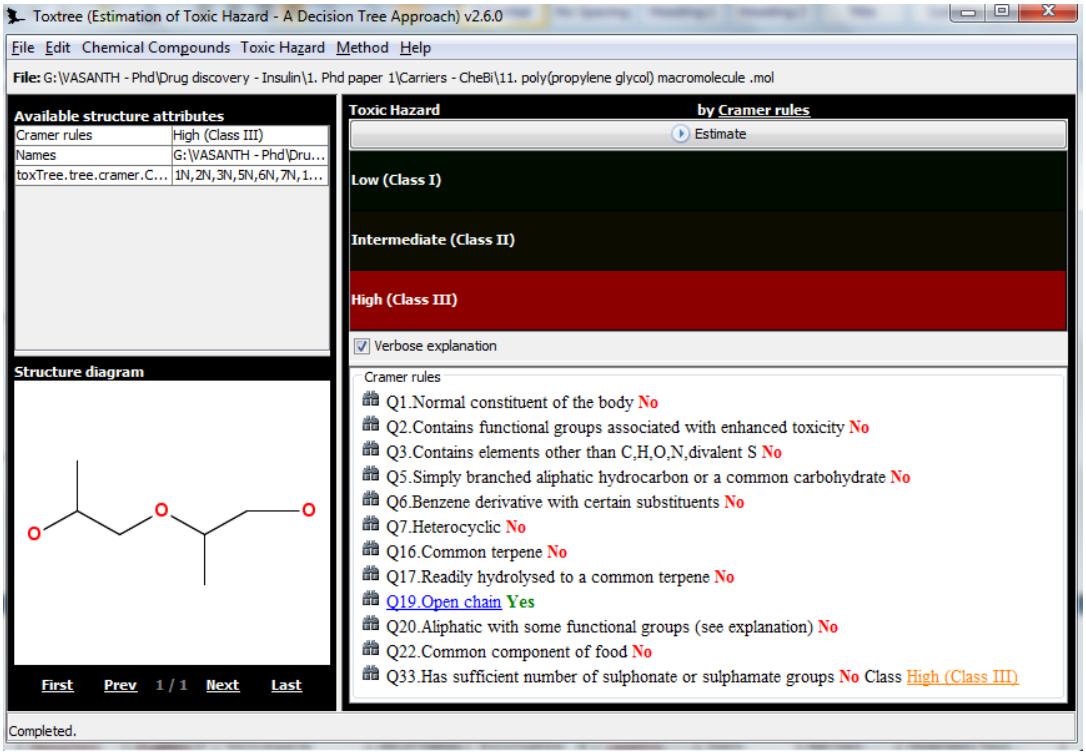
Supplementary figure 4(i)- C9 – Chitosan



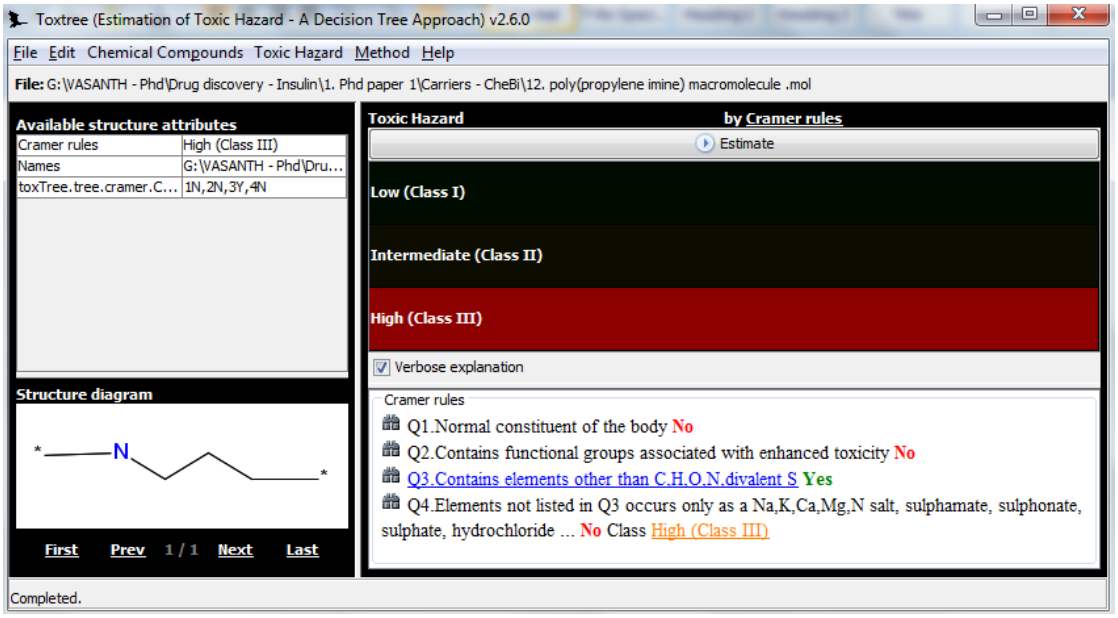
Supplementary figure 4(j)- C10 – Pectin



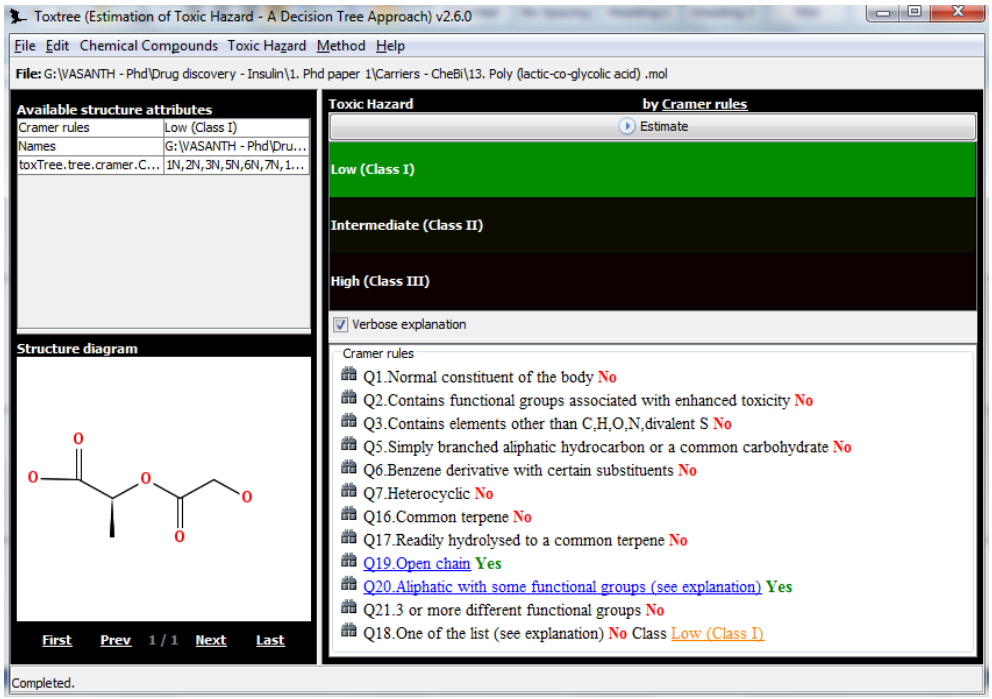
Supplementary figure 4(k)- C11 - Poly(propylene glycol)



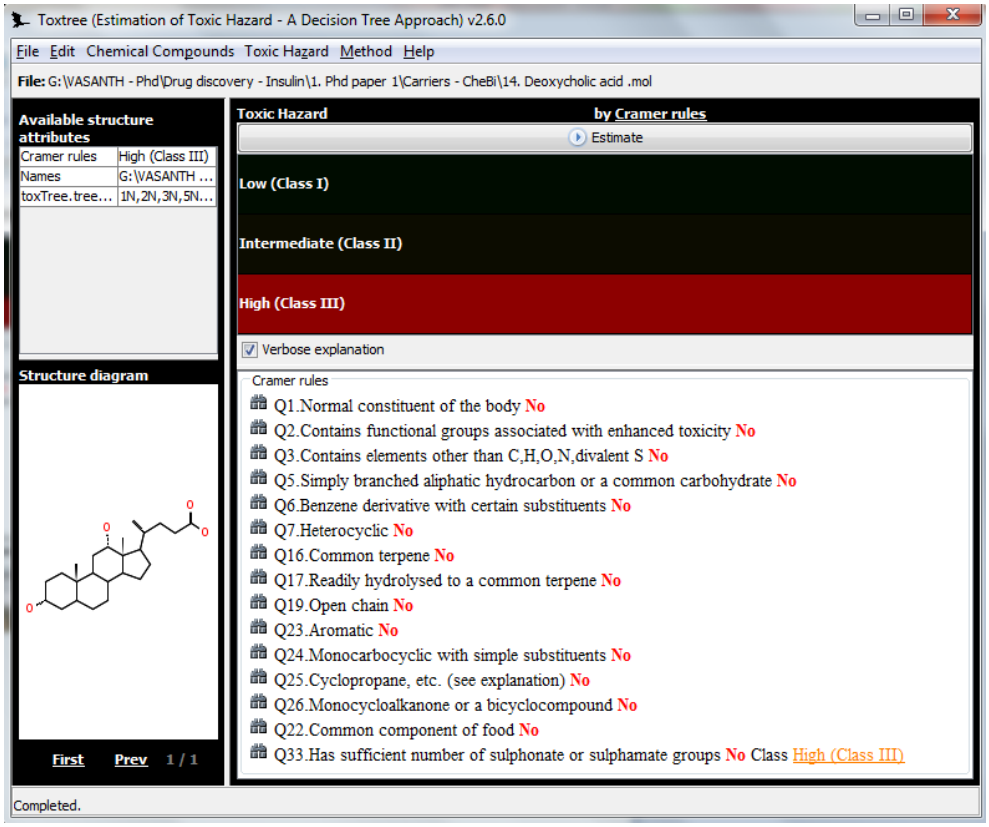
Supplementary figure 4(l)- C12 - Poly(propylene imine)



Supplementary figure 4(m)- C13 - Poly (lactic-co-glycolic acid)



Supplementary figure 4(n)- C14 - Deoxycholic acid

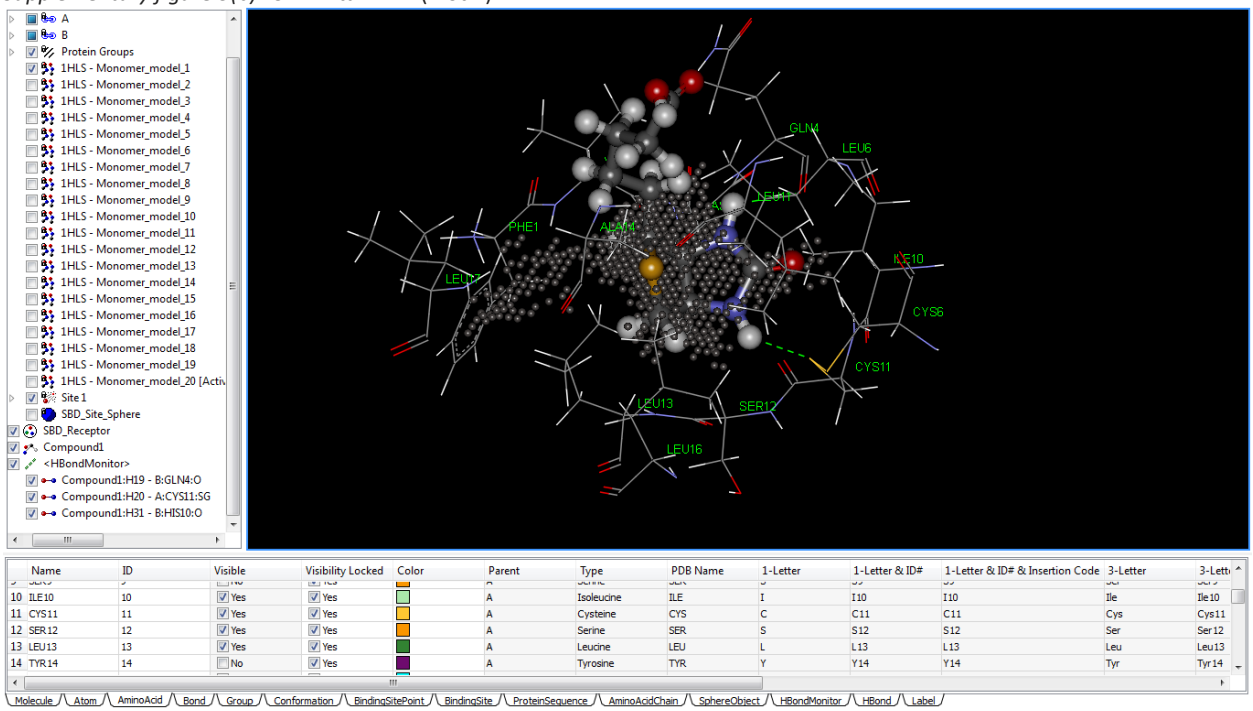


Supplementary figure 5

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)

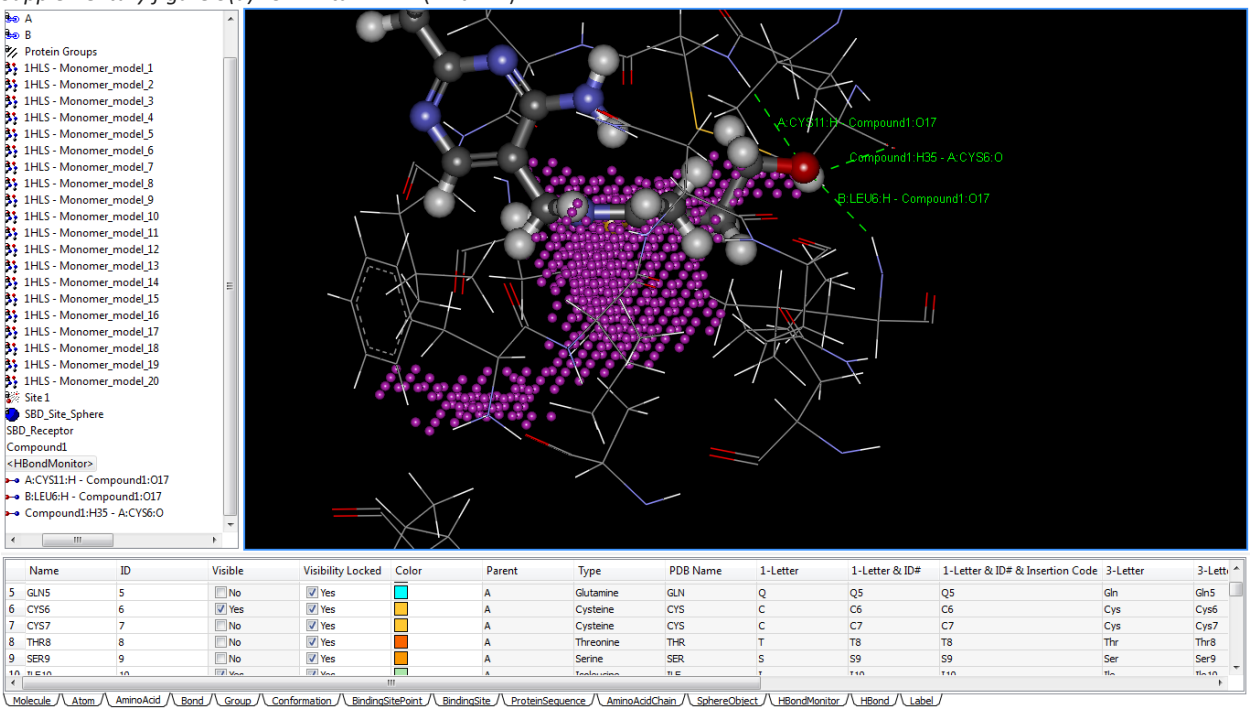


	Name	ID	Visible	Visibility Locked	Color	Parent	Type	PDB Name	1-Letter	1-Letter & ID#	1-Letter & ID# & Insertion Code	3-Letter	3-Lett
10	ILE10	10	Yes	Yes		A	Isoleucine	ILE	I	I10	I10	Ile	Ile10
11	CYS11	11	Yes	Yes		A	Cysteine	CYS	C	C11	C11	Cys	Cys11
12	SER12	12	Yes	Yes		A	Serine	SER	S	S12	S12	Ser	Ser12
13	LEU13	13	Yes	Yes		A	Leucine	LEU	L	L13	L13	Leu	Leu13
14	TYR14	14	No	Yes		A	Tyrosine	TYR	Y	Y14	Y14	Tyr	Tyr14

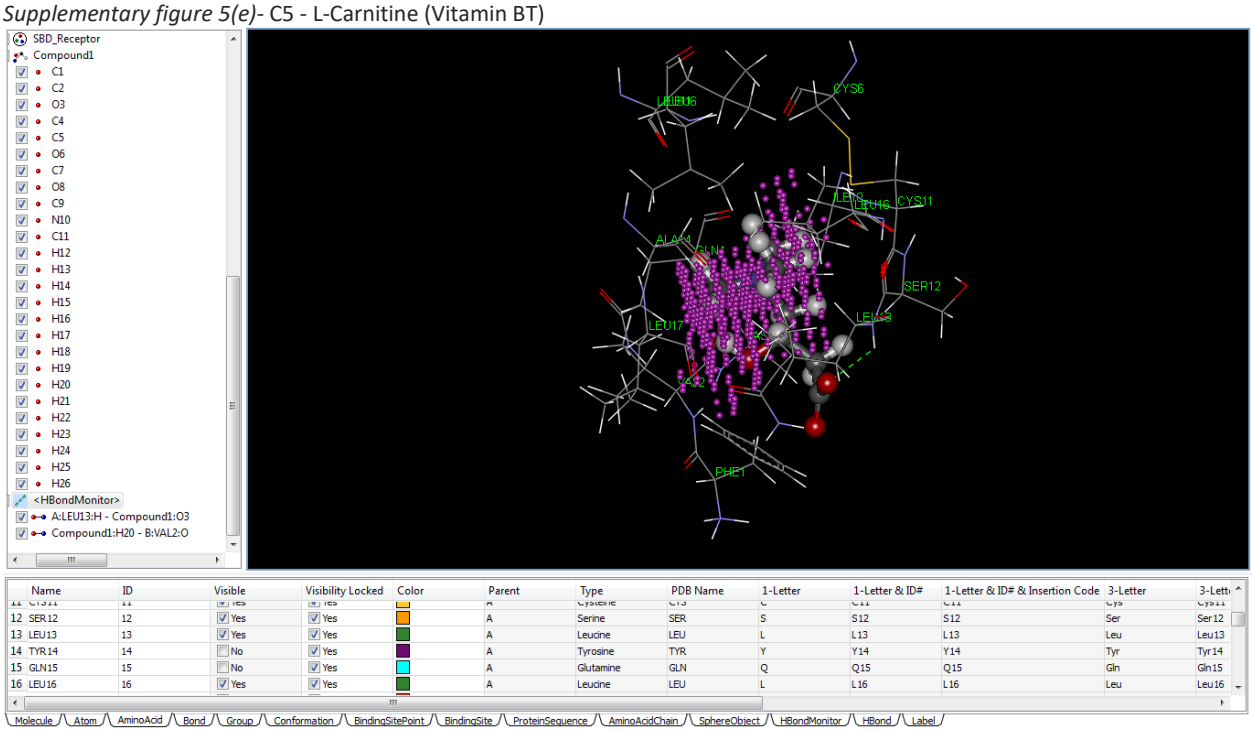
Molecule | Atom | AminoAcid | Bond | Group | Conformation | BindingSitePoint | BindingSite | ProteinSequence | AminoAcidChain | SphereObject | HBondMonitor | HBond | Label

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

Supplementary figure 5(d)- C4 - Vitamin B1 (Thiamin)

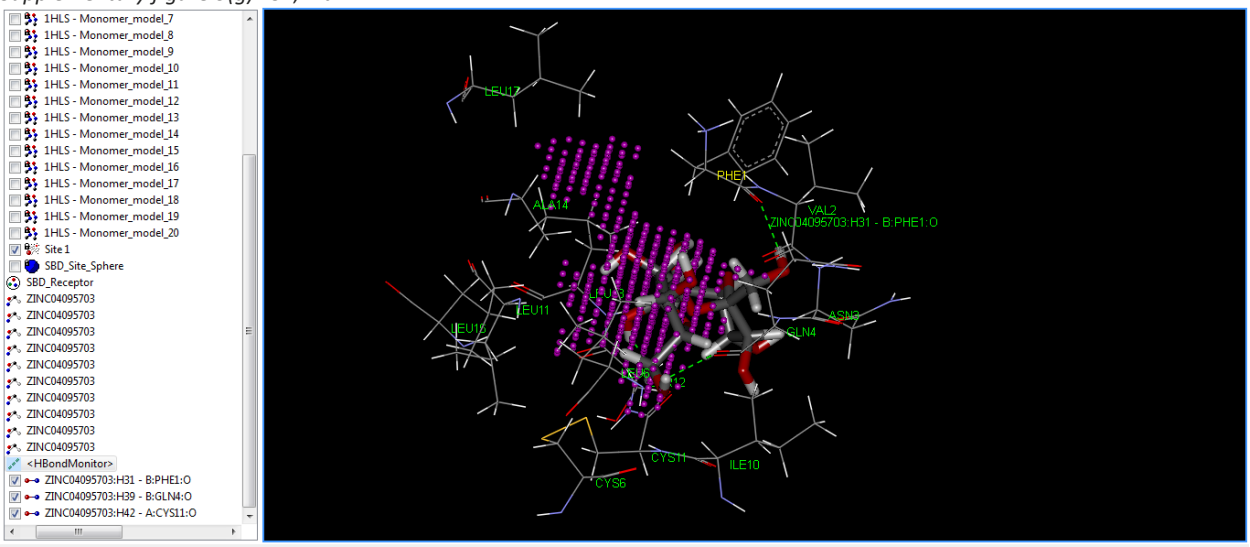


Supplementary figure 5(e)- C5 - L-Carnitine (Vitamin BT)



Supplementary figure 5(f)- C6 - Poly-N-vinylpyrrolidone
No Conjugation

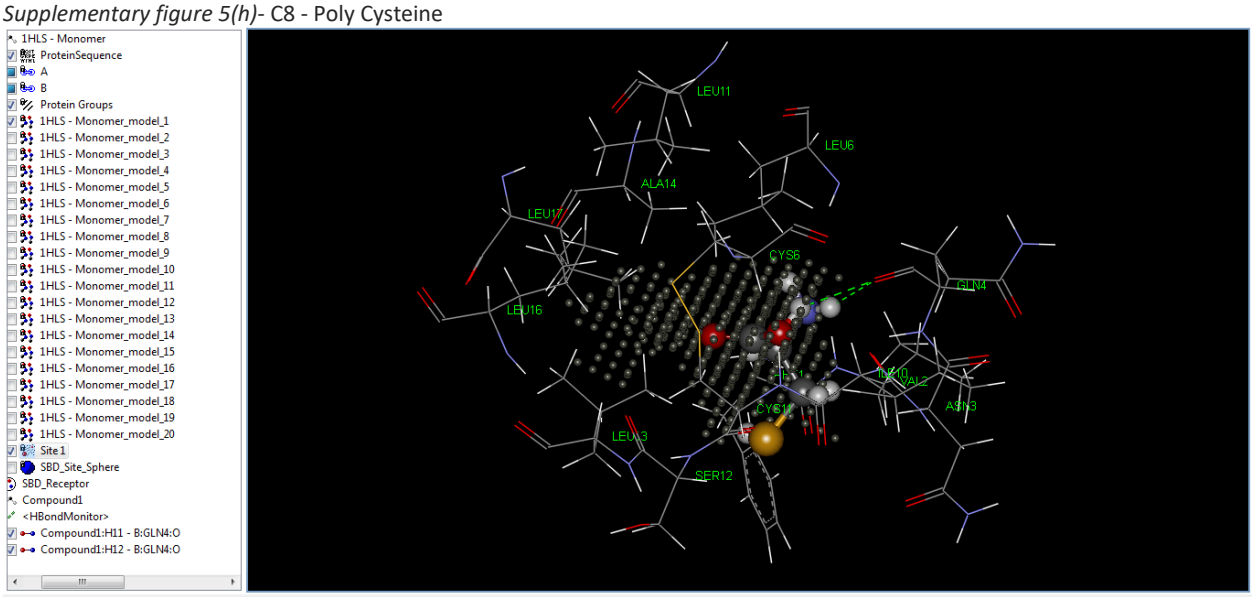
Supplementary figure 5(g)- C7; Inulin



	Name	ID	Visible	Visibility Locked	Color	Parent	Type	PDB Name	1-Letter	1-Letter & ID#	1-Letter & ID# & Insertion Code	3-Letter	3-Letter
21	ASN21	21	No	Yes		A	Asparagine	ASN	N	N21	N21	Asn	Asn21
22	PHE1	1	Yes	Yes		B	Phenylalanine	PHE	F	F1	F1	Phe	Phe1
23	VAL2	2	Yes	Yes		B	Valine	VAL	V	V2	V2	Val	Val2
24	ASN3	3	Yes	Yes		B	Asparagine	ASN	N	N3	N3	Asn	Asn3
25	GLN4	4	Yes	Yes		B	Glutamine	GLN	Q	Q4	Q4	Gln	Gln4
26	HIS5	5	No	Yes		B	Histidine	HIS	H	H5	H5	His	His5

Molecule | Atom | AminoAcid | Bond | Group | Conformation | BindingSitePoint | BindingSite | ProteinSequence | AminoAcidChain | SphereObject | HBondMonitor | HBond | Label

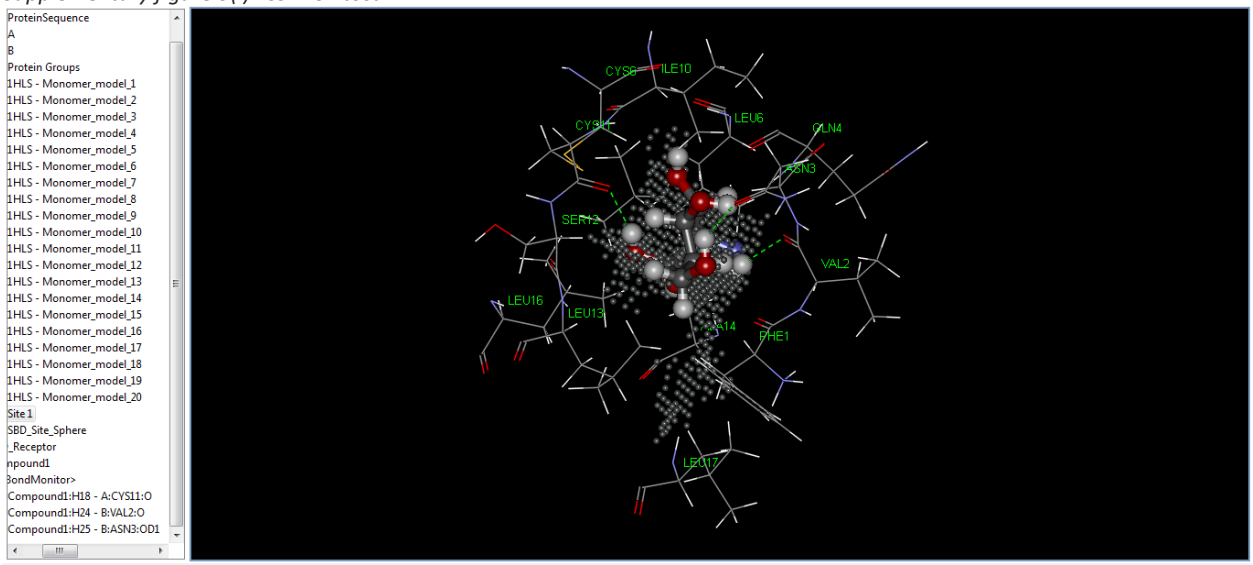
Supplementary figure 5(h)- C8 - Poly Cysteine



	Name	ID	Visible	Visibility Locked	Color	Parent	Type	PDB Name	1-Letter	1-Letter & ID#	1-Letter & ID# & Insertion Code	3-Letter	3-Letter
24	ASN3	3	Yes	Yes		B	Asparagine	ASN	N	N3	N3	Asn	Asn3
25	GLN4	4	Yes	Yes		B	Glutamine	GLN	Q	Q4	Q4	Gln	Gln4
26	HIS5	5	No	Yes		B	Histidine	HIS	H	H5	H5	His	His5
27	LEU6	6	Yes	Yes		B	Leucine	LEU	L	L6	L6	Leu	Leu6

Molecule | Atom | AminoAcid | Bond | Group | Conformation | BindingSitePoint | BindingSite | ProteinSequence | AminoAcidChain | SphereObject | HBondMonitor | HBond | Label

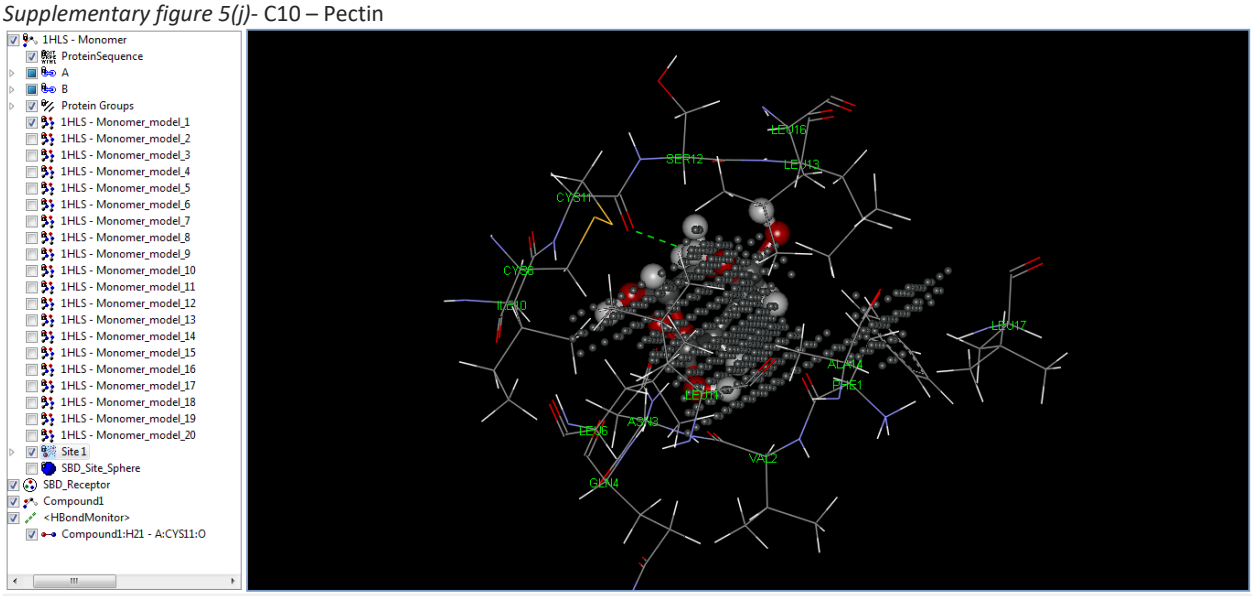
Supplementary figure 5(i)- C9 – Chitosan



	Name	ID	Visible	Visibility Locked	Color	Parent	Type	PDB Name	1-Letter	1-Letter & ID#	1-Letter & ID# & Insertion Code	3-Letter	3-Lett
22	PHE1	1	Yes	Yes		B	Phenylalanine	PHE	F	F1	F1	Phe	Phe1
23	VAL2	2	Yes	Yes		B	Valine	VAL	V	V2	V2	Val	Val2
24	ASN3	3	Yes	Yes		B	Asparagine	ASN	N	N3	N3	Asn	Asn3
25	GLN4	4	Yes	Yes		B	Glutamine	GLN	Q	Q4	Q4	Gln	Gln4
26	HIS5	5	No	Yes		B	Histidine	HIS	H	H5	H5	His	His5

Molecule | Atom | AminoAcid | Bond | Group | Conformation | BindingSitePoint | BindingSite | ProteinSequence | AminoAcidChain | SphereObject | HBondMonitor | HBond | Label

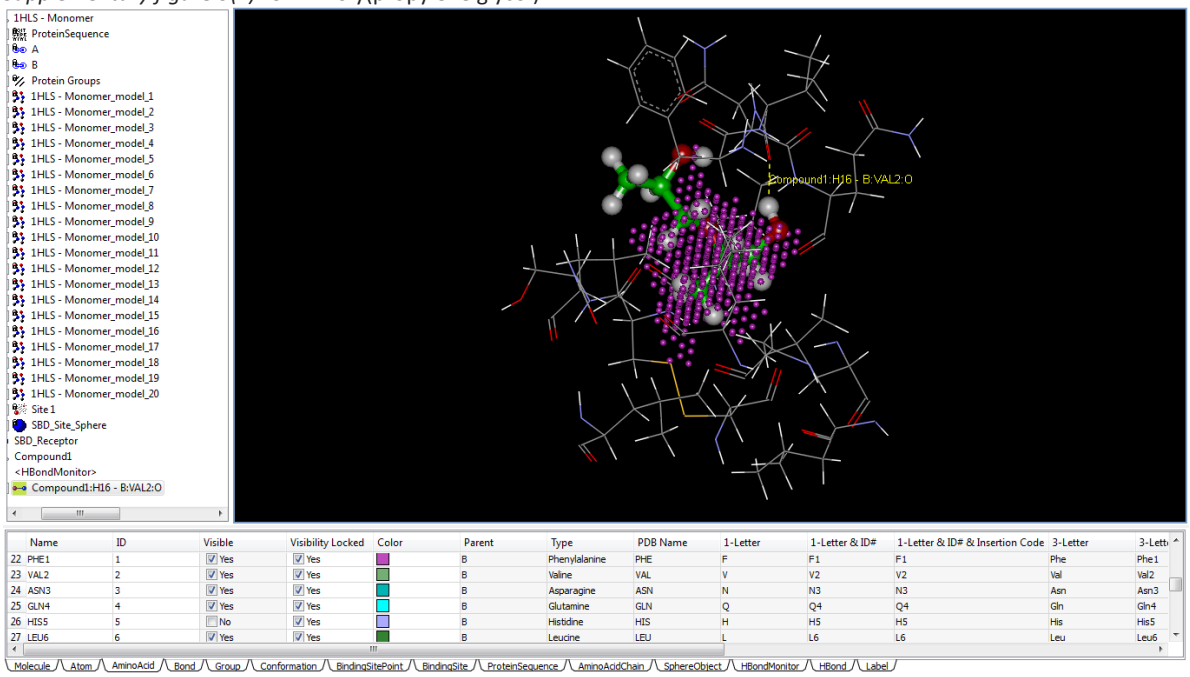
Supplementary figure 5(j)- C10 – Pectin



	Name	ID	Visible	Visibility Locked	Color	Parent	Type	PDB Name	1-Letter	1-Letter & ID#	1-Letter & ID# & Insertion Code	3-Letter	3-Lett
10	ILE10	10	Yes	Yes		A	Isoleucine	ILE	I	I10	I10	Ile	Ile10
11	CYS11	11	Yes	Yes		A	Cysteine	CYS	C	C11	C11	Cys	Cys11
12	SER12	12	Yes	Yes		A	Serine	SER	S	S12	S12	Ser	Ser12
13	LEU13	13	Yes	Yes		A	Leucine	LEU	L	L13	L13	Leu	Leu13

Molecule | Atom | AminoAcid | Bond | Group | Conformation | BindingSitePoint | BindingSite | ProteinSequence | AminoAcidChain | SphereObject | HBondMonitor | HBond | Label

Supplementary figure 5(k)- C11 - Poly(propylene glycol)

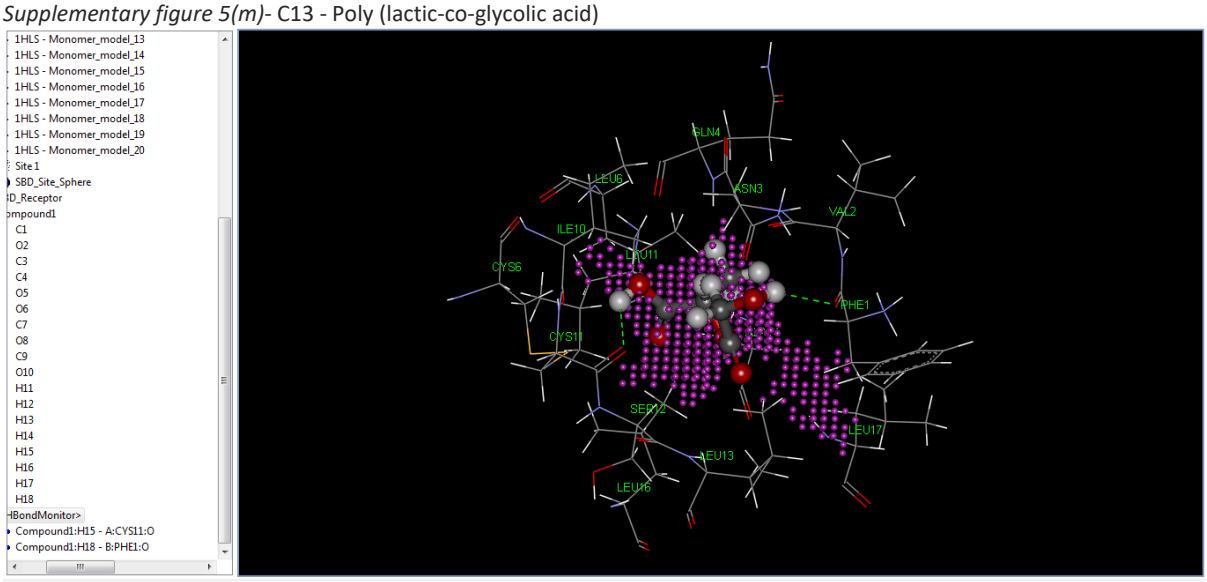


	Name	ID	Visible	Visibility Locked	Color	Parent	Type	PDB Name	1-Letter	1-Letter & ID#	1-Letter & ID# & Insertion Code	3-Letter	3-Letti
22	PHE1	1	Yes	Yes		B	Phenylalanine	PHE	F	F1	F1	Phe	Phe1
23	VAL2	2	Yes	Yes		B	Valine	VAL	V	V2	V2	Val	Val2
24	ASN3	3	Yes	Yes		B	Asparagine	ASN	N	N3	N3	Asn	Asn3
25	GLN4	4	Yes	Yes		B	Glutamine	GLN	Q	Q4	Q4	Gln	Gln4
26	HIS5	5	No	Yes		B	Histidine	HIS	H	H5	H5	His	His5
27	LEU6	6	Yes	Yes		B	Leucine	LEU	L	L6	L6	Leu	Leu6

Molecule | Atom | AminoAcid | Bond | Group | Conformation | BindingSitePoint | BindingSite | ProteinSequence | AminoAcidChain | SphereObject | HBondMonitor | HBond | Label

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

Supplementary figure 5(m)- C13 - Poly (lactic-co-glycolic acid)



	Name	ID	Visible	Visibility Locked	Color	Parent	Type	PDB Name	1-Letter	1-Letter & ID#	1-Letter & ID# & Insertion Code	3-Letter	3-Letti
21	ASN21	21	No	Yes		A	Asparagine	ASN	N	N21	N21	Asn	Asn21
22	PHE1	1	Yes	Yes		B	Phenylalanine	PHE	F	F1	F1	Phe	Phe1
23	VAL2	2	Yes	Yes		B	Valine	VAL	V	V2	V2	Val	Val2
24	ASN3	3	Yes	Yes		B	Asparagine	ASN	N	N3	N3	Asn	Asn3

Molecule | Atom | AminoAcid | Bond | Group | Conformation | BindingSitePoint | BindingSite | ProteinSequence | AminoAcidChain | SphereObject | HBondMonitor | HBond | Label

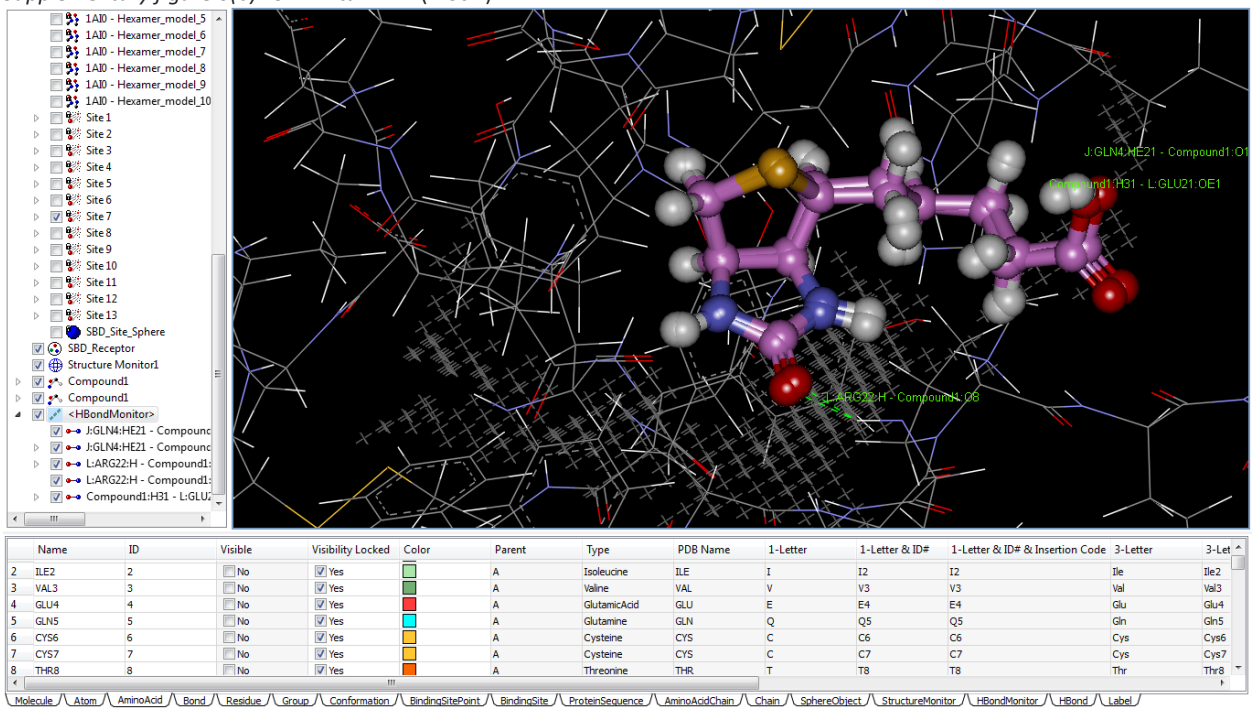
Supplementary figure 5(n)- C14 - Deoxycholic acid
No Conjugation

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

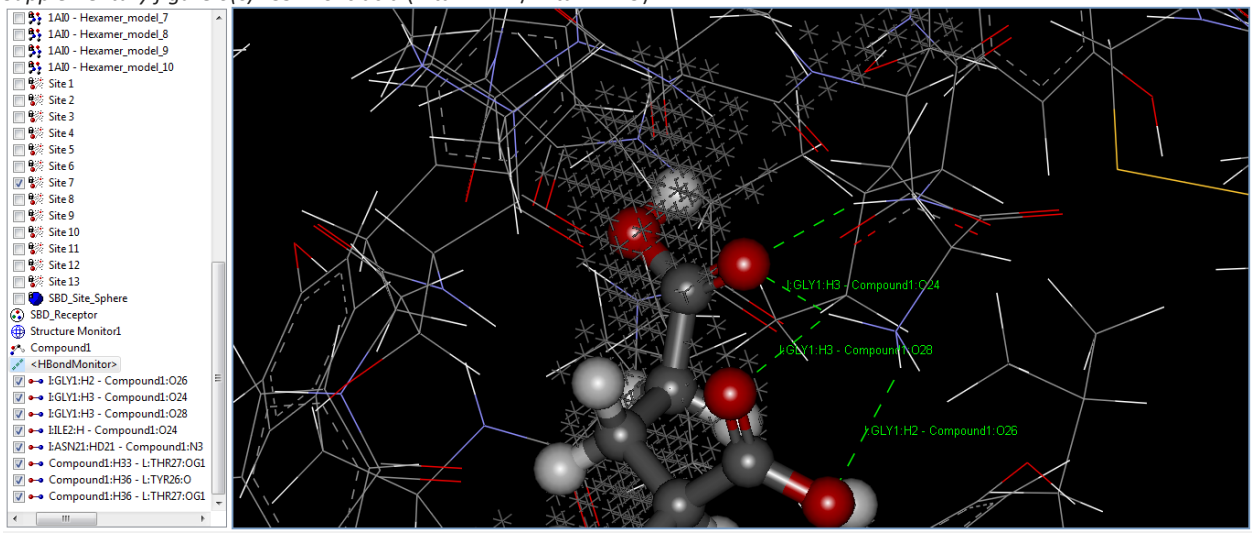
Supplementary figure 6(b)- C2 - Vitamin H (Biotin)



	Name	ID	Visible	Visibility Locked	Color	Parent	Type	PDB Name	1-Letter	1-Letter & ID#	1-Letter & ID# & Insertion Code	3-Letter	3-Let
2	ILE2	2	No	Yes		A	Isoleucine	ILE	I	I2	I2	Ile	Ile2
3	VAL3	3	No	Yes		A	Valine	VAL	V	V3	V3	Val	Val3
4	GLU4	4	No	Yes		A	GlutamicAcid	GLU	E	E4	E4	Glu	Glu4
5	GLN5	5	No	Yes		A	Glutamine	GLN	Q	Q5	Q5	Gln	Gln5
6	CYS6	6	No	Yes		A	Cysteine	CYS	C	C6	C6	Cys	Cys6
7	CYS7	7	No	Yes		A	Cysteine	CYS	C	C7	C7	Cys	Cys7
8	THR8	8	No	Yes		A	Threonine	THR	T	T8	T8	Thr	Thr8

Molecule | Atom | AminoAcid | Bond | Residue | Group | Conformation | BindingSitePoint | BindingSite | ProteinSequence | AminoAcidChain | Chain | SphereObject | StructureMonitor | HBondMonitor | HBond | Label

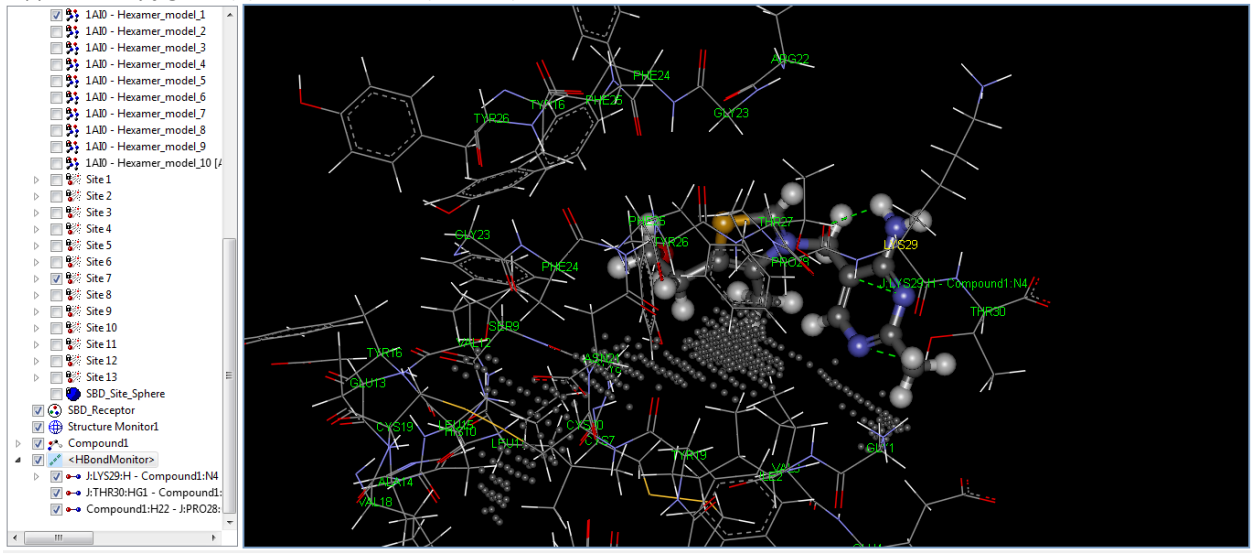
Supplementary figure 6(c)- C3 - Folic acid (Vitamin M / Vitamin B9)



	Name	ID	Visible	Visibility Locked	Color	Parent	Type	PDB Name	1-Letter	1-Letter & ID#	1-Letter & ID# & Insertion Code	3-Letter	3-Let
204	THR30	30	No	Yes		H	Threonine	THR	T	T30	T30	Thr	Thr30
205	GLY1	1	Yes	Yes		I	Glycine	GLY	G	G1	G1	Gly	Gly1
206	ILE2	2	Yes	Yes		I	Isoleucine	ILE	I	I2	I2	Ile	Ile2
207	VAL3	3	Yes	Yes		I	Valine	VAL	V	V3	V3	Val	Val3
208	GLU4	4	Yes	Yes		I	GlutamicAcid	GLU	E	E4	E4	Glu	Glu4
209	GLN5	5	Yes	Yes		I	Glutamine	GLN	Q	Q5	Q5	Gln	Gln5
210	CYS6	6	Yes	Yes		I	Cysteine	CYS	C	C6	C6	Cys	Cys6

Molecule | Atom | AminoAcid | Bond | Residue | Group | Conformation | BindingSitePoint | BindingSite | ProteinSequence | AminoAcidChain | Chain | SphereObject | StructureMonitor | HBondMonitor | HBond | Label

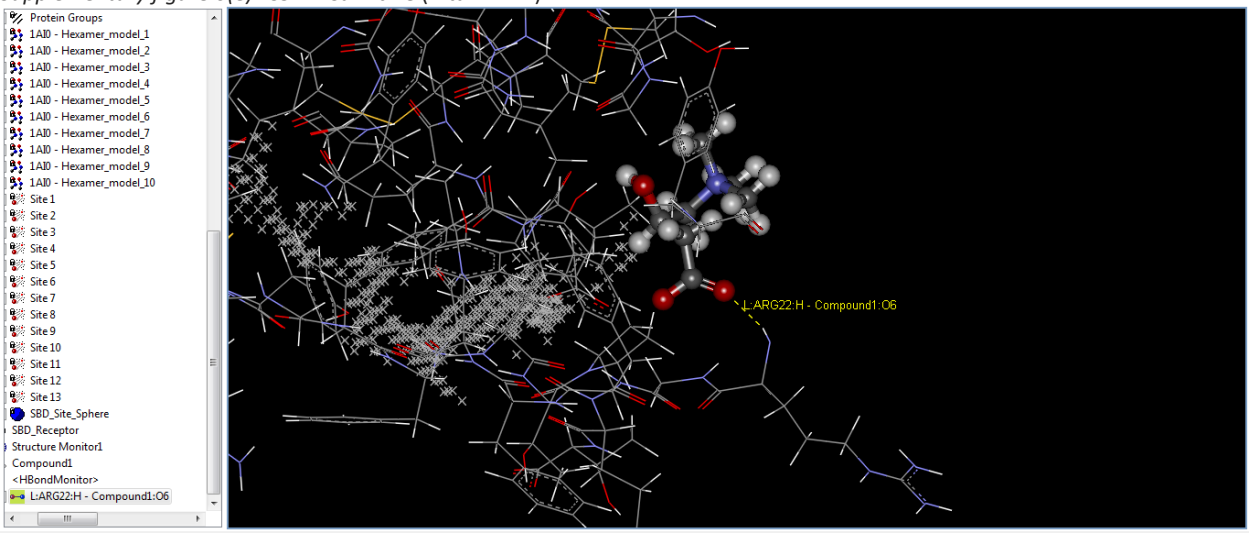
Supplementary figure 6(d)- C4 - Vitamin B1 (Thiamin)



	Name	ID	Visible	Visibility Locked	Color	Parent	Type	PDB Name	1-Letter	1-Letter & ID#	1-Letter & ID# & Insertion Code	3-Letter	3-Let
253	PRO28	28	Yes	Yes		J	Proline	PRO	P	P28	P28	Pro	Pro28
254	LYS29	29	Yes	Yes		J	Lysine	LYS	K	K29	K29	Lys	Lys29
255	THR30	30	Yes	Yes		J	Threonine	THR	T	T30	T30	Thr	Thr30
256	GLY1	1	No	Yes		K	Glycine	GLY	G	G1	G1	Gly	Gly1
257	ILE2	2	No	Yes		K	Isoleucine	ILE	I	I2	I2	Ile	Ile2

Molecule | Atom | AminoAcid | Bond | Residue | Group | Conformation | BindingSitePoint | BindingSite | ProteinSequence | AminoAcidChain | Chain | SphereObject | StructureMonitor | HBondMonitor | HBond | Label

Supplementary figure 6(e)- C5 - L-Carnitine (Vitamin BT)

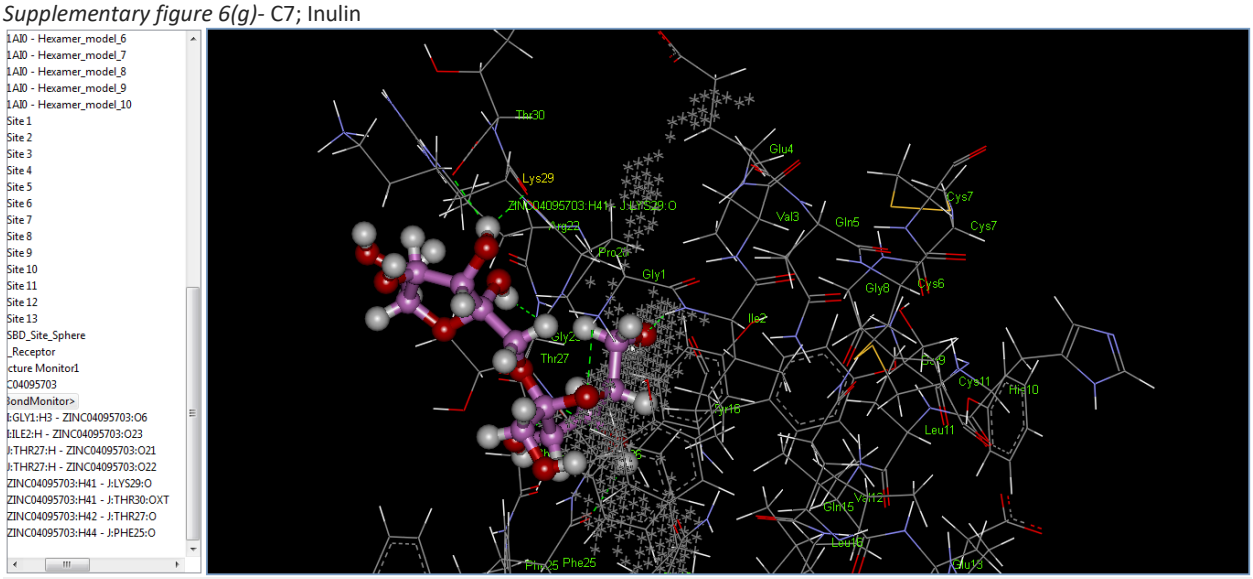


	Name	ID	Visible	Visibility Locked	Color	Parent	Type	PDB Name	1-Letter	1-Letter & ID#	1-Letter & ID# & Insertion Code	3-Letter	3-Let
297	GLU21	21	No	Yes		L	GlutamicAcid	GLU	E	E21	E21	Glu	Glu21
298	ARG22	22	Yes	Yes		L	Arginine	ARG	R	R22	R22	Arg	Arg22
299	GLY23	23	Yes	Yes		L	Glycine	GLY	G	G23	G23	Gly	Gly23
300	PHE24	24	Yes	Yes		L	Phenylalanine	PHE	F	F24	F24	Phe	Phe24
301	PHE25	25	Yes	Yes		L	Phenylalanine	PHE	F	F25	F25	Phe	Phe25
302	TYR26	26	Yes	Yes		L	Tyrosine	TYR	Y	Y26	Y26	Tyr	Tyr26
303	THR27	27	No	Yes		L	Threonine	THR	T	T27	T27	Thr	Thr27

Molecule | Atom | AminoAcid | Bond | Residue | Group | Conformation | BindingSitePoint | BindingSite | ProteinSequence | AminoAcidChain | Chain | SphereObject | StructureMonitor | HBondMonitor | HBond | Label

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

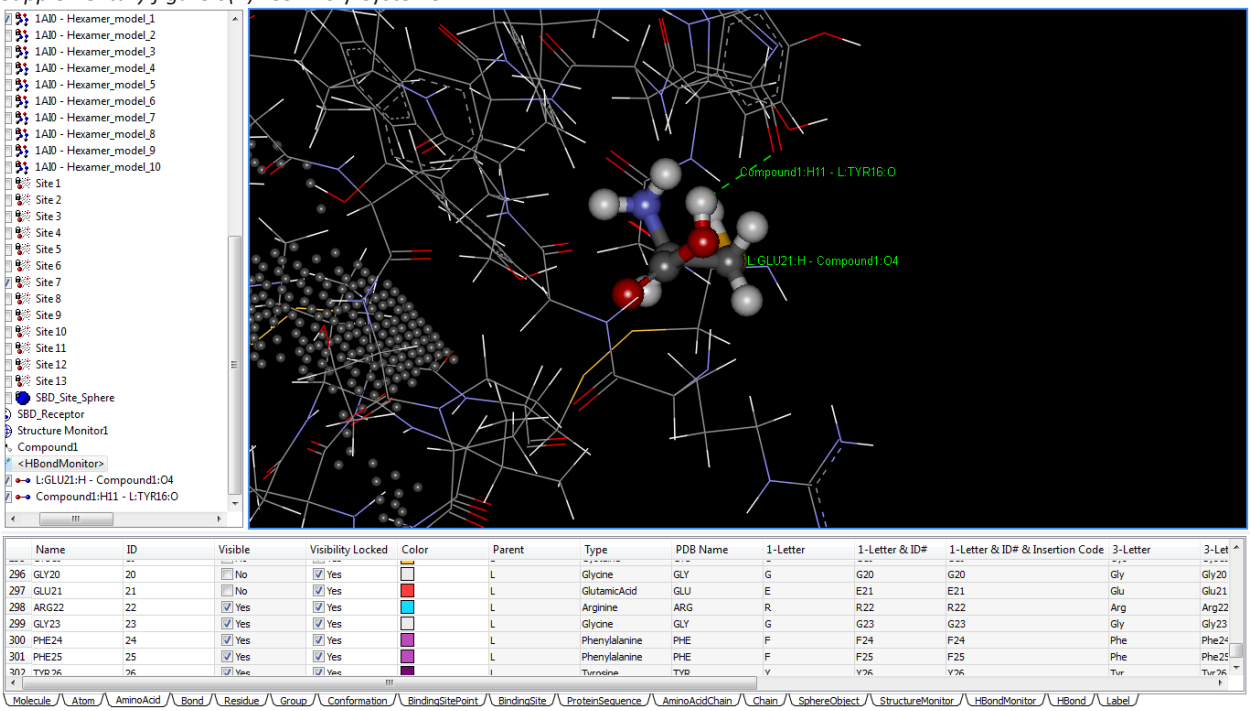
Supplementary figure 6(g)- C7; Inulin



	Name	ID	Visible	Visibility Locked	Color	Parent	Type	PDB Name	1-Letter	1-Letter & ID#	1-Letter & ID# & Insertion Code	3-Letter	3-Let
253	PRO28	28	Yes	Yes		J	Proline	PRO	P	P28	P28	Pro	Pro28
254	LYS29	29	Yes	Yes		J	Lysine	LYS	K	K29	K29	Lys	Lys29
255	THR30	30	Yes	Yes		J	Threonine	THR	T	T30	T30	Thr	Thr30
256	GLY1	1	No	Yes		K	Glycine	GLY	G	G1	G1	Gly	Gly1
257	ILE2	2	No	Yes		K	Isoleucine	ILE	I	I2	I2	Ile	Ile2

Molecule | Atom | AminoAcid | Bond | Residue | Group | Conformation | BindingSitePoint | BindingSite | ProteinSequence | AminoAcidChain | Chain | SphereObject | StructureMonitor | HBondMonitor | HBond | Label

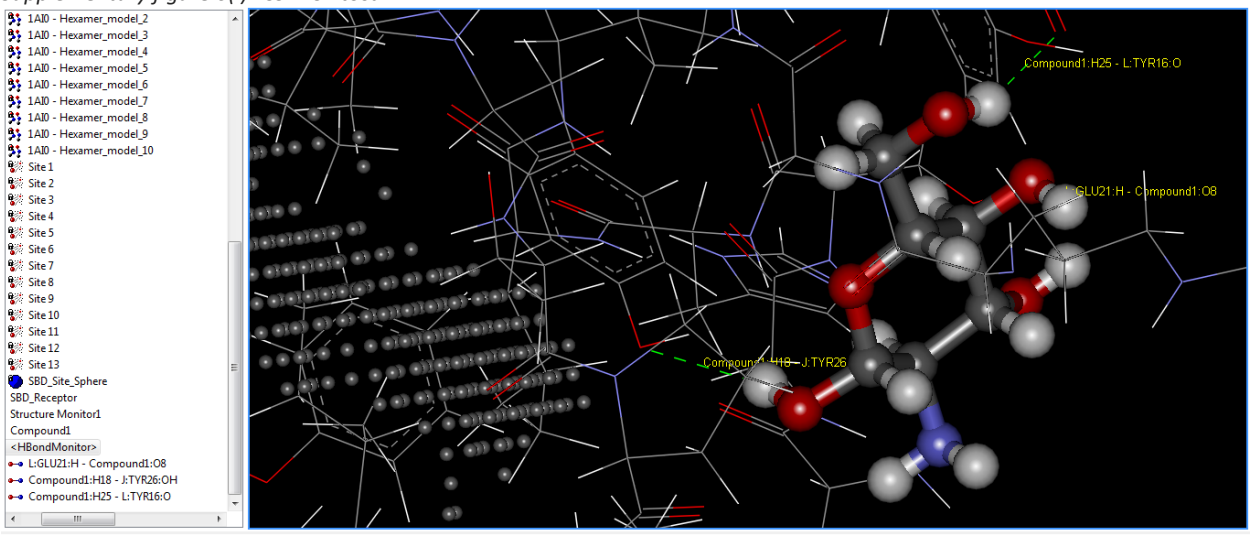
Supplementary figure 6(h)- C8 - Poly Cysteine



	Name	ID	Visible	Visibility Locked	Color	Parent	Type	PDB Name	1-Letter	1-Letter & ID#	1-Letter & ID# & Insertion Code	3-Letter	3-Let
296	GLY20	20	No	Yes		L	Glycine	GLY	G	G20	G20	Gly	Gly20
297	GLU21	21	No	Yes		L	GlutamicAcid	GLU	E	E21	E21	Glu	Glu21
298	ARG22	22	Yes	Yes		L	Arginine	ARG	R	R22	R22	Arg	Arg22
299	GLY23	23	Yes	Yes		L	Glycine	GLY	G	G23	G23	Gly	Gly23
300	PHE24	24	Yes	Yes		L	Phenylalanine	PHE	F	F24	F24	Phe	Phe24
301	PHE25	25	Yes	Yes		L	Phenylalanine	PHE	F	F25	F25	Phe	Phe25
302	TYR26	26	Yes	Yes		L	Tyrosine	TYR	Y	Y26	Y26	Tyr	Tyr26

Molecule | Atom | AminoAcid | Bond | Residue | Group | Conformation | BindingSitePoint | BindingSite | ProteinSequence | AminoAcidChain | Chain | SphereObject | StructureMonitor | HBondMonitor | HBond | Label

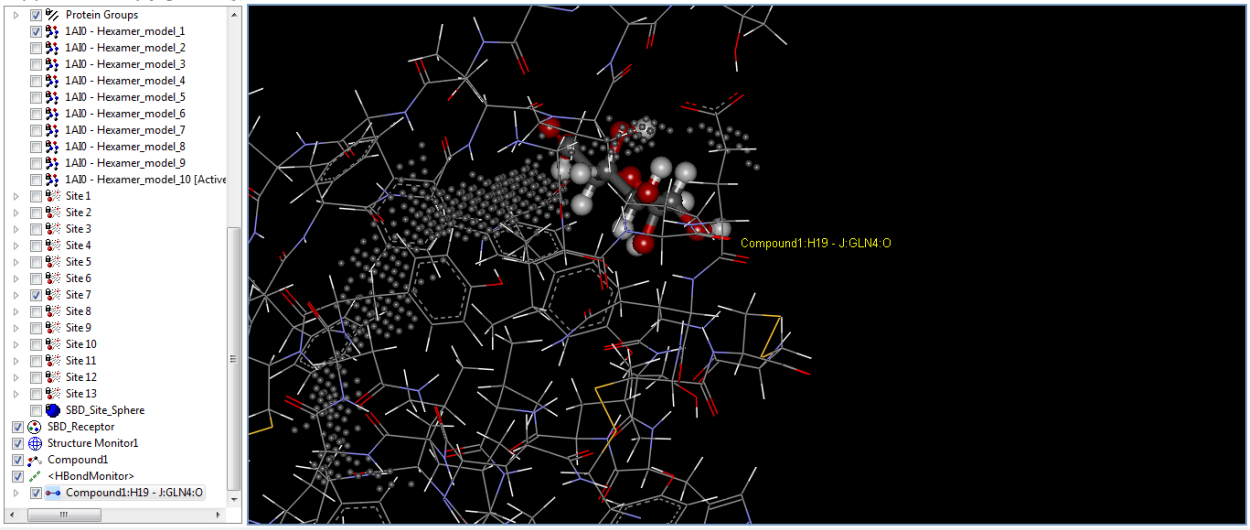
Supplementary figure 6(i)- C9 – Chitosan



	Name	ID	Visible	Visibility Locked	Color	Parent	Type	PDB Name	1-Letter	1-Letter & ID#	1-Letter & ID# & Insertion Code	3-Letter	3-Let
239	ALA14	14	Yes	Yes		J	Alanine	ALA	A	A14	A14	Ala	Ala14
240	LEU15	15	Yes	Yes		J	Leucine	LEU	L	L15	L15	Leu	Leu15
241	TYR16	16	Yes	Yes		J	Tyrosine	TYR	Y	Y16	Y16	Tyr	Tyr16
242	LEU17	17	No	Yes		J	Leucine	LEU	L	L17	L17	Leu	Leu17
243	VAL18	18	Yes	Yes		J	Valine	VAL	V	V18	V18	Val	Val18
244	CYS19	19	Yes	Yes		J	Cysteine	CYS	C	C19	C19	Cys	Cys19
245	GLY20	20	No	Yes		J	Glycine	GLY	G	G20	G20	Gly	Gly20

Molecule | Atom | AminoAcid | Bond | Residue | Group | Conformation | BindingSitePoint | BindingSite | ProteinSequence | AminoAcidChain | Chain | SphereObject | StructureMonitor | HBondMonitor | HBond | Label

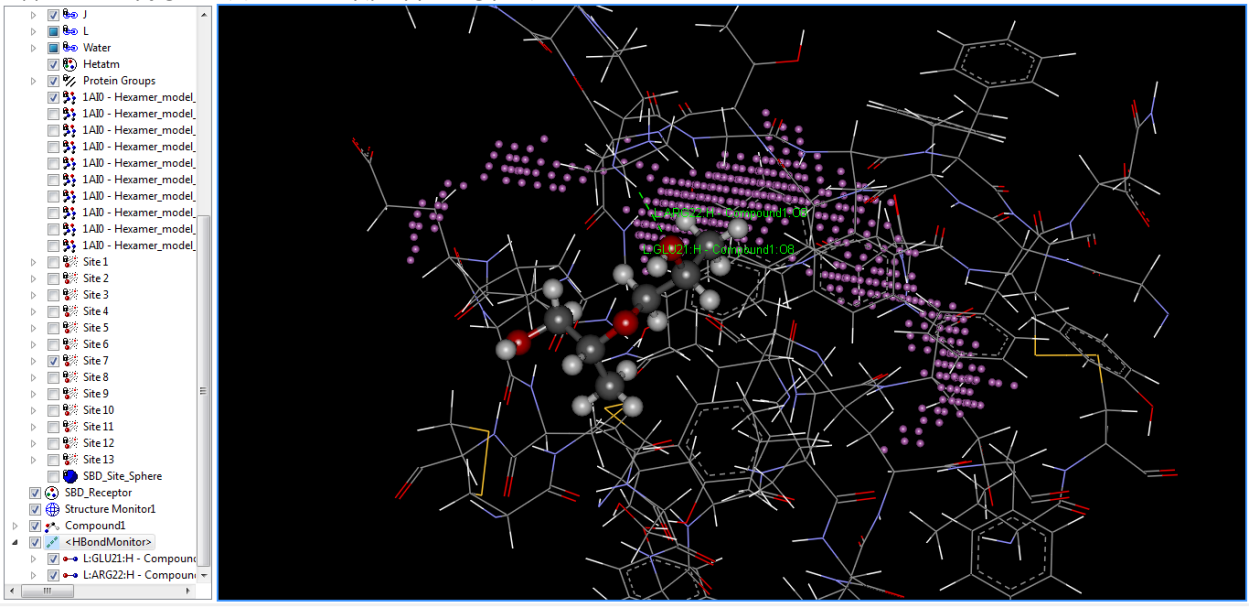
Supplementary figure 6(j)- C10 – Pectin



	Name	ID	Visible	Visibility Locked	Color	Parent	Type	PDB Name	1-Letter	1-Letter & ID#	1-Letter & ID# & Insertion Code	3-Letter	3-Let
228	ASN3	3	No	Yes		J	Asparagine	ASN	N	N3	N3	Asn	Asn3
229	GLN4	4	No	Yes		J	Glutamine	GLN	Q	Q4	Q4	Gln	Gln4
230	HIS5	5	No	Yes		J	Histidine	HIS	H	H5	H5	His	His5
231	LEU6	6	No	Yes		J	Leucine	LEU	L	L6	L6	Leu	Leu6
232	CYS7	7	Yes	Yes		J	Cysteine	CYS	C	C7	C7	Cys	Cys7
233	GLY8	8	Yes	Yes		J	Glycine	GLY	G	G8	G8	Gly	Gly8

Molecule | Atom | AminoAcid | Bond | Residue | Group | Conformation | BindingSitePoint | BindingSite | ProteinSequence | AminoAcidChain | Chain | SphereObject | StructureMonitor | HBondMonitor | HBond | Label

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

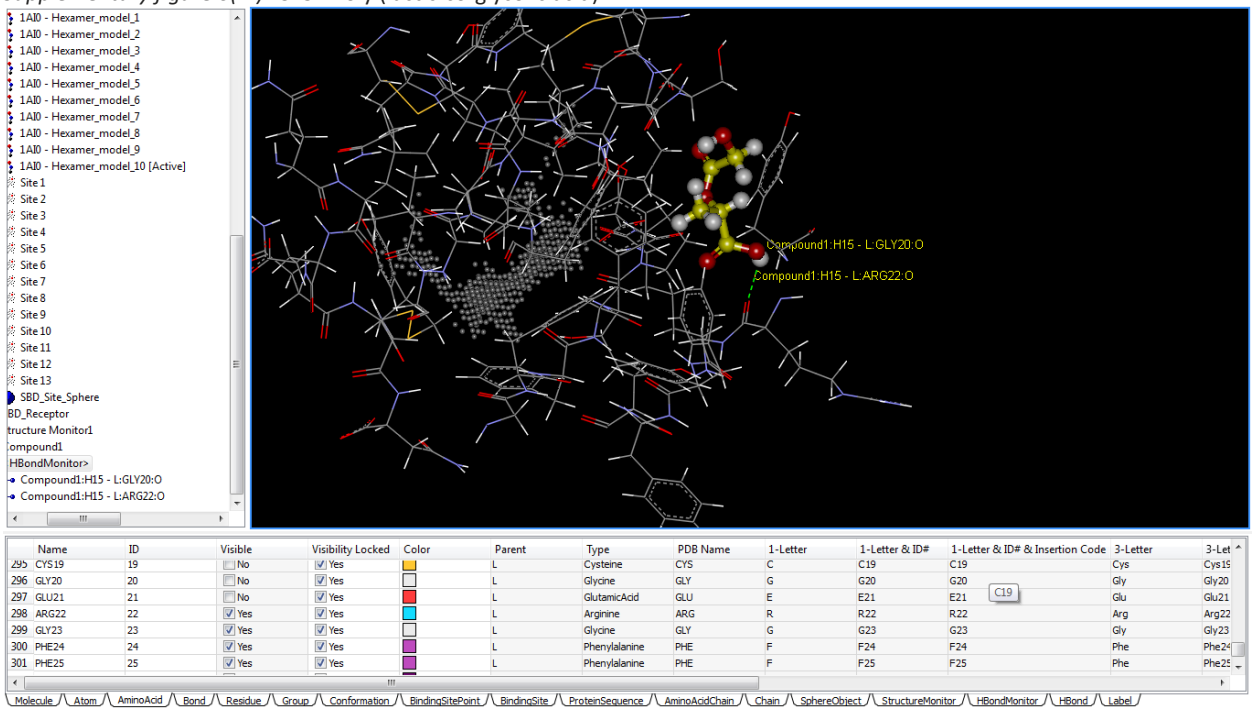


	Name	ID	Visible	Visibility Locked	Color	Parent	Type	PDB Name	1-Letter	1-Letter & ID#	1-Letter & ID# & Insertion Code	3-Letter	3-Let
297	GLU21	21	No	Yes		L	GlutamicAcid	GLU	E	E21	E21	Glu	Glu21
298	ARG22	22	Yes	Yes		L	Arginine	ARG	R	R22	R22	Arg	Arg22

Molecule | Atom | AminoAcid | Bond | Residue | Group | Conformation | BindingSitePoint | BindingSite | ProteinSequence | AminoAcidChain | Chain | SphereObject | StructureMonitor | HBondMonitor | HBond | Label

Supplementary figure 6(l)- C12 - Poly(propylene imine)
No Conjugation

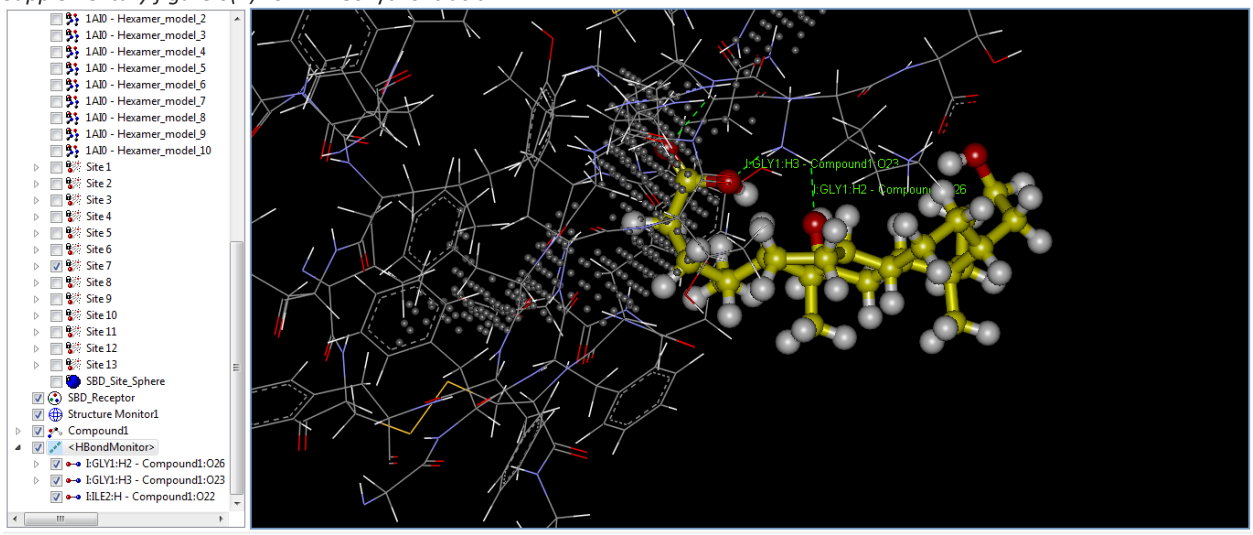
Supplementary figure 6(m)- C13 - Poly (lactic-co-glycolic acid)



	Name	ID	Visible	Visibility Locked	Color	Parent	Type	PDB Name	1-Letter	1-Letter & ID#	1-Letter & ID# & Insertion Code	3-Letter	3-Let
295	CYS19	19	No	Yes		L	Cysteine	CYS	C	C19	C19	Cys	Cys19
296	GLY20	20	No	Yes		L	Glycine	GLY	G	G20	G20	Gly	Gly20
297	GLU21	21	No	Yes		L	GlutamicAcid	GLU	E	E21	E21	Glu	Glu21
298	ARG22	22	Yes	Yes		L	Arginine	ARG	R	R22	R22	Arg	Arg22
299	GLY23	23	Yes	Yes		L	Glycine	GLY	G	G23	G23	Gly	Gly23
300	PHE24	24	Yes	Yes		L	Phenylalanine	PHE	F	F24	F24	Phe	Phe24
301	PHE25	25	Yes	Yes		L	Phenylalanine	PHE	F	F25	F25	Phe	Phe25

Molecule | Atom | AminoAcid | Bond | Residue | Group | Conformation | BindingSitePoint | BindingSite | ProteinSequence | AminoAcidChain | Chain | SphereObject | StructureMonitor | HBondMonitor | HBond | Label

Supplementary figure 6(n)- C14 - Deoxycholic acid



	Name	ID	Visible	Visibility Locked	Color	Parent	Type	PDB Name	1-Letter	1-Letter & ID#	1-Letter & ID# & Insertion Code	3-Letter	3-Let
202	PRO28	28	No	Yes		H	Proline	PRO	P	P28	P28	Pro	Pro28
203	LYS29	29	No	Yes		H	Lysine	LYS	K	K29	K29	Lys	Lys29
204	THR30	30	No	Yes		H	Threonine	THR	T	T30	T30	Thr	Thr30
205	GLY1	1	Yes	Yes		I	Glycine	GLY	G	G1	G1	Gly	Gly1
206	ILE2	2	Yes	Yes		I	Isoleucine	ILE	I	I2	I2	Ile	Ile2
207	VAL3	3	Yes	Yes		I	Valine	VAL	V	V3	V3	Val	Val3
208	GLU4	4	Yes	Yes		I	GlutamicAcid	GLU	E	E4	E4	Glu	Glu4

Molecule | Atom | AminoAcid | Bond | Residue | Group | Conformation | BindingSitePoint | BindingSite | ProteinSequence | AminoAcidChain | Chain | SphereObject | StructureMonitor | HBondMonitor | HBond | Label

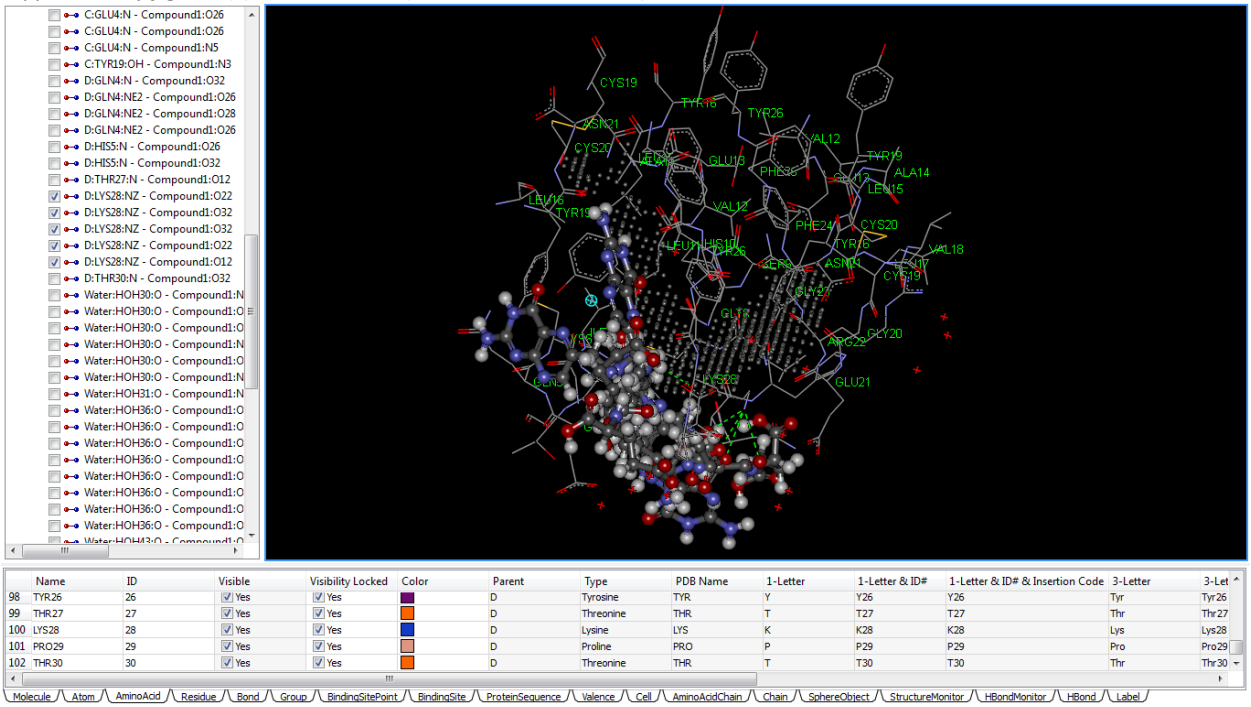
Supplementary figure 7

Conjugation results of Insulin Lispro (PDB ID: 1 LPH), 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

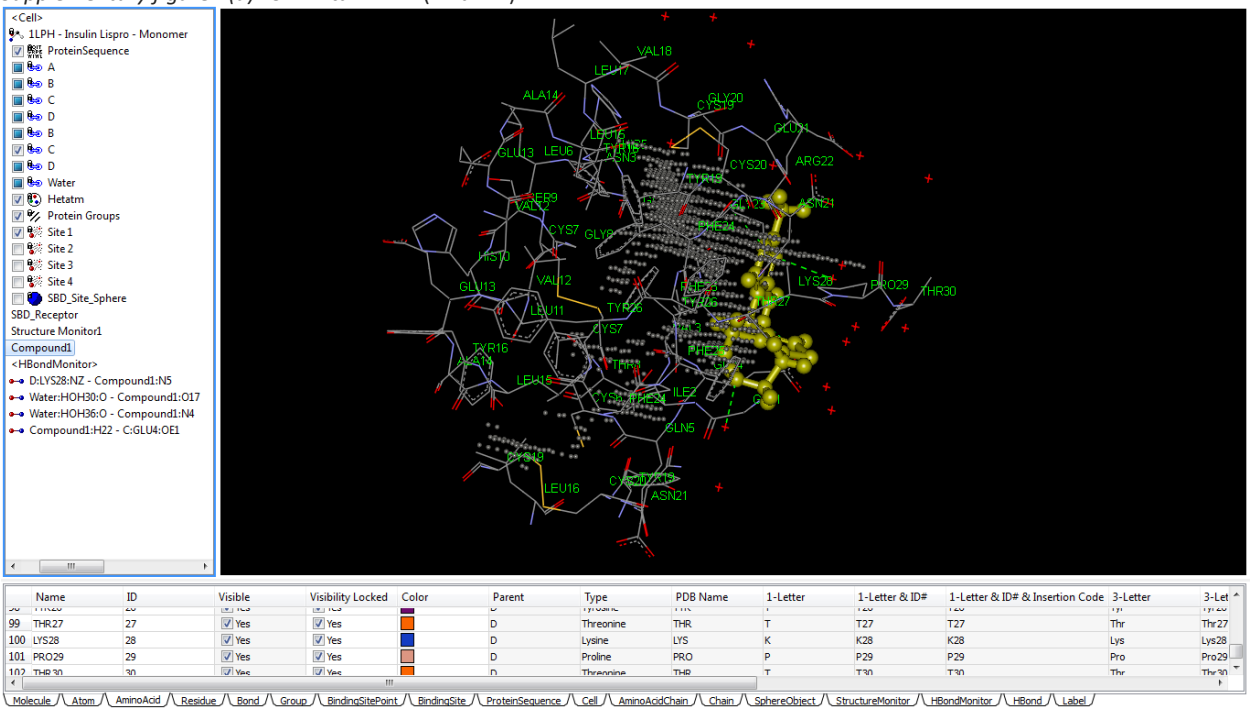
Supplementary figure 7(c)- C3 - Folic acid (Vitamin M / Vitamin B9)



	Name	ID	Visible	Visibility Locked	Color	Parent	Type	PDB Name	1-Letter	1-Letter & ID#	1-Letter & ID# & Insertion Code	3-Letter	3-Let
98	TYR26	26	Yes	Yes		D	Tyrosine	TYR	Y	Y26	Y26	Tyr	Tyr26
99	THR27	27	Yes	Yes		D	Threonine	THR	T	T27	T27	Thr	Thr27
100	LYS28	28	Yes	Yes		D	Lysine	LYS	K	K28	K28	Lys	Lys28
101	PRO29	29	Yes	Yes		D	Proline	PRO	P	P29	P29	Pro	Pro29
102	THR30	30	Yes	Yes		D	Threonine	THR	T	T30	T30	Thr	Thr30

Molecule | Atom | AminoAcid | Residue | Bond | Group | BindingSitePoint | BindingSite | ProteinSequence | Valence | Cell | AminoAcidChain | Chain | SphereObject | StructureMonitor | HBondMonitor | HBond | Label

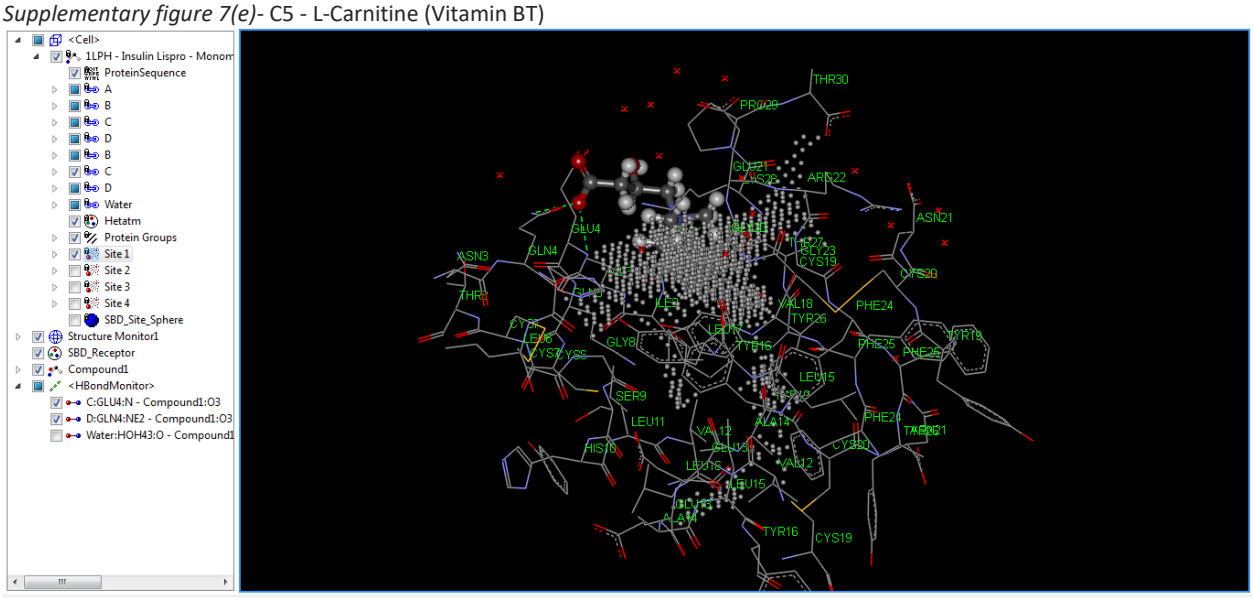
Supplementary figure 7(d)- C4 - Vitamin B1 (Thiamin)



	Name	ID	Visible	Visibility Locked	Color	Parent	Type	PDB Name	1-Letter	1-Letter & ID#	1-Letter & ID# & Insertion Code	3-Letter	3-Let
99	THR27	27	Yes	Yes		D	Threonine	THR	T	T27	T27	Thr	Thr27
100	LYS28	28	Yes	Yes		D	Lysine	LYS	K	K28	K28	Lys	Lys28
101	PRO29	29	Yes	Yes		D	Proline	PRO	P	P29	P29	Pro	Pro29

Molecule | Atom | AminoAcid | Residue | Bond | Group | BindingSitePoint | BindingSite | ProteinSequence | Cell | AminoAcidChain | Chain | SphereObject | StructureMonitor | HBondMonitor | HBond | Label

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

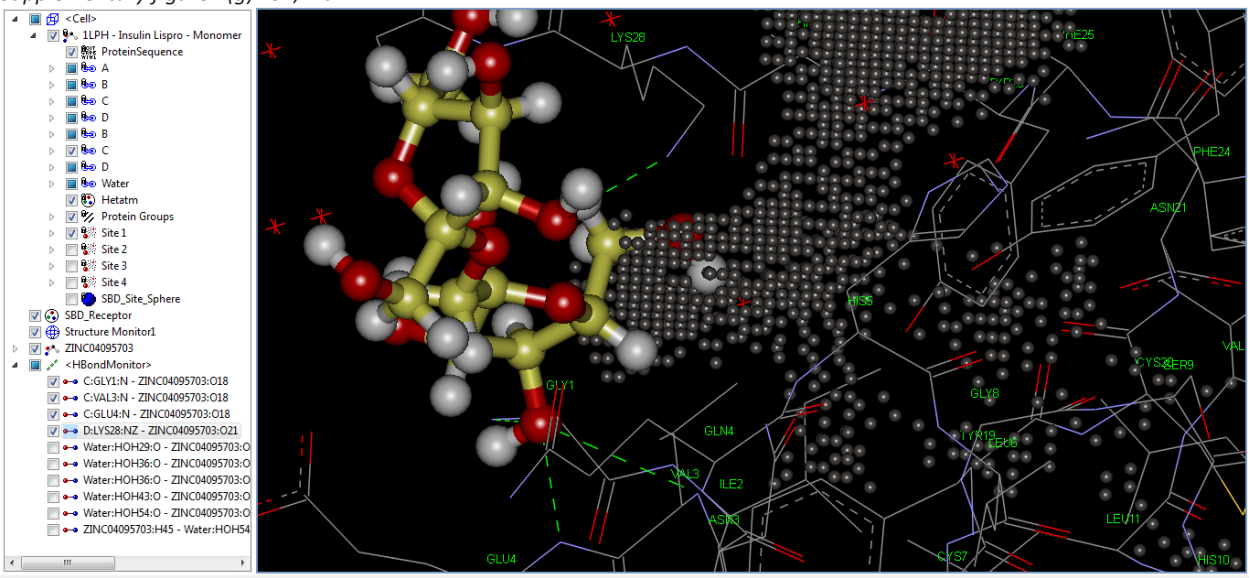


	Name	ID	Visible	Visibility Locked	Color	Parent	Type	PDB Name	1-Letter	1-Letter & ID#	1-Letter & ID# & Insertion Code	3-Letter	3-Let
54	VAL3	3	Yes	Yes		C	Valine	VAL	V	V3	V3	Val	Val3
55	GLU4	4	Yes	Yes		C	GlutamicAcid	GLU	E	E4	E4	Glu	Glu4
56	GLN5	5	Yes	Yes		C	Glutamine	GLN	Q	Q5	Q5	Gln	Gln5
57	CYS6	6	Yes	Yes		C	Cysteine	CYS	C	C6	C6	Cys	Cys6

Molecule | Atom | AminoAcid | Residue | Bond | Group | BindingSitePoint | BindingSite | ProteinSequence | Valence | Cell | AminoAcidChain | Chain | SphereObject | StructureMonitor | Label | HBondMonitor | HBond

Supplementary figure 7(f)- C6 - Poly-N-vinylpyrrolidone
No Conjugation

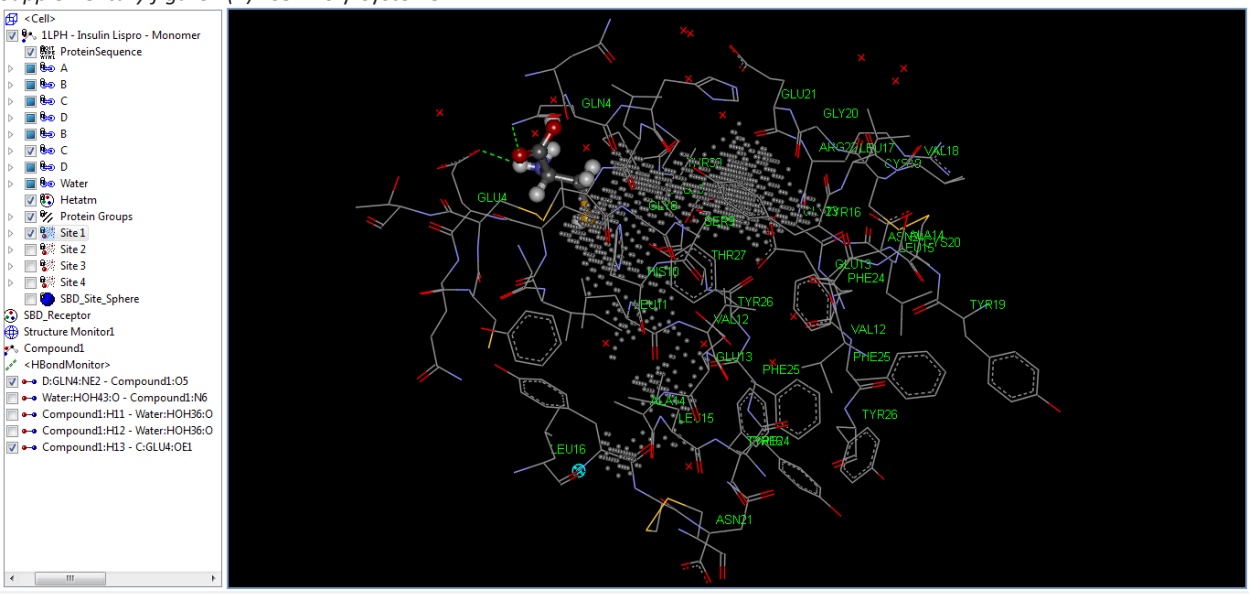
Supplementary figure 7(g)- C7; Inulin



	Name	ID	Visible	Visibility Locked	Color	Parent	Type	PDB Name	1-Letter	1-Letter & ID#	1-Letter & ID# & Insertion Code	3-Letter	3-Let
99	THR27	27	Yes	Yes		D	Threonine	THR	T	T27	T27	Thr	Thr27
100	LYS28	28	Yes	Yes		D	Lysine	LYS	K	K28	K28	Lys	Lys28
101	PRO29	29	Yes	Yes		D	Proline	PRO	P	P29	P29	Pro	Pro29
102	THR30	30	Yes	Yes		D	Threonine	THR	T	T30	T30	Thr	Thr30

Molecule | Atom | AminoAcid | Residue | Bond | Group | BindingSitePoint | BindingSite | ProteinSequence | Cell | AminoAcidChain | Chain | SphereObject | StructureMonitor | HBondMonitor | HBond | Label

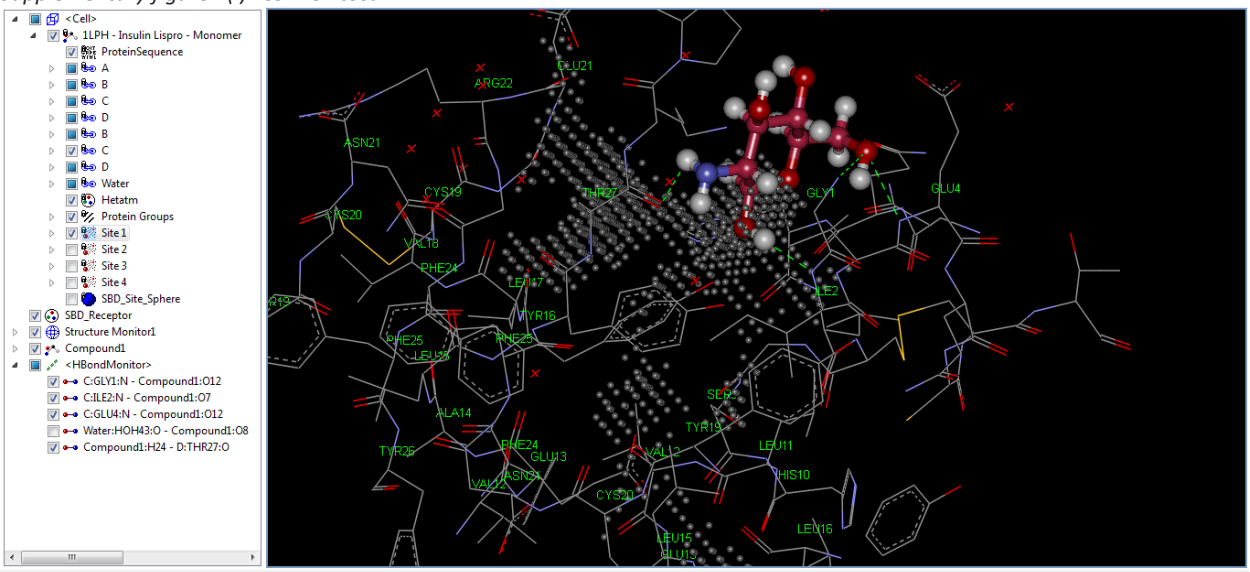
Supplementary figure 7(h)- C8 - Poly Cysteine



	Name	ID	Visible	Visibility Locked	Color	Parent	Type	PDB Name	1-Letter	1-Letter & ID#	1-Letter & ID# & Insertion Code	3-Letter	3-Let
54	VAL3	3	Yes	Yes		C	Valine	VAL	V	V3	V3	Val	Val3
55	GLU4	4	Yes	Yes		C	GlutamicAcid	GLU	E	E4	E4	Glu	Glu4
56	GLN5	5	Yes	Yes		C	Glutamine	GLN	Q	Q5	Q5	Gln	Gln5

Molecule | Atom | AminoAcid | Residue | Bond | Group | BindingSitePoint | BindingSite | ProteinSequence | Valence | Cell | AminoAcidChain | Chain | SphereObject | StructureMonitor | HBondMonitor | HBond | Label

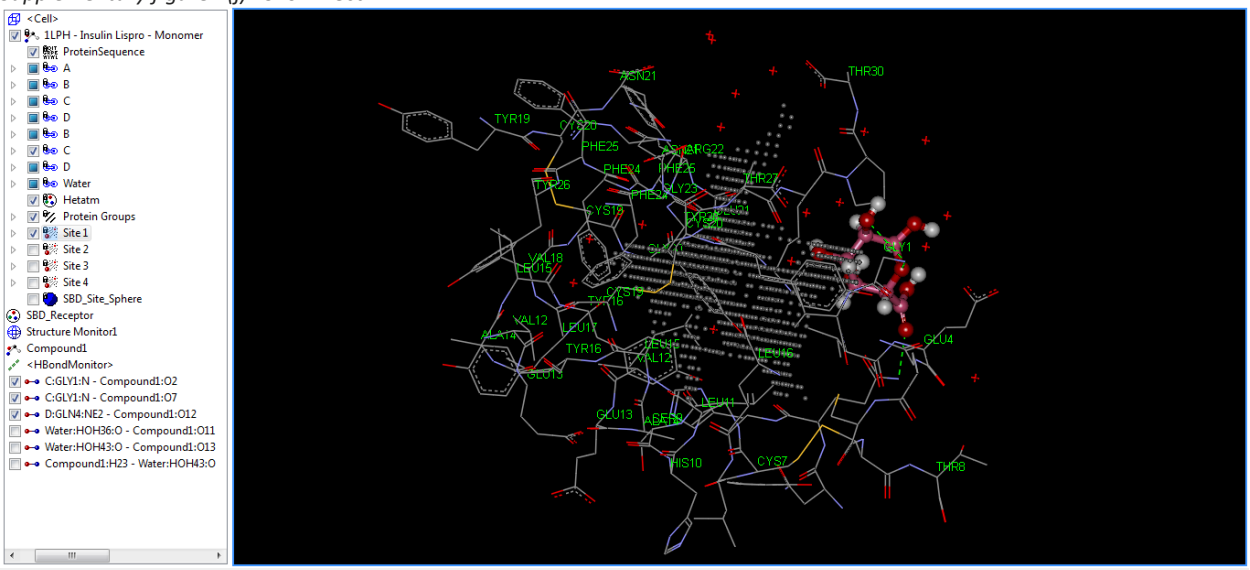
Supplementary figure 7(i)- C9 – Chitosan



	Name	ID	Visible	Visibility Locked	Color	Parent	Type	PDB Name	1-Letter	1-Letter & ID#	1-Letter & ID# & Insertion Code	3-Letter	3-Let
52	GLY1	1	Yes	Yes		C	Glycine	GLY	G	G1	G1	Gly	Gly1
53	ILE2	2	Yes	Yes		C	Isoleucine	ILE	I	I2	I2	Ile	Ile2
54	VAL3	3	Yes	Yes		C	Valine	VAL	V	V3	V3	Val	Val3
55	GLU4	4	Yes	Yes		C	GlutamicAcid	GLU	E	E4	E4	Glu	Glu4

Molecule | Atom | AminoAcid | Residue | Bond | Group | BindingSitePoint | BindingSite | ProteinSequence | Valence | Cell | AminoAcidChain | Chain | SphereObject | StructureMonitor | HBondMonitor | HBond | Label

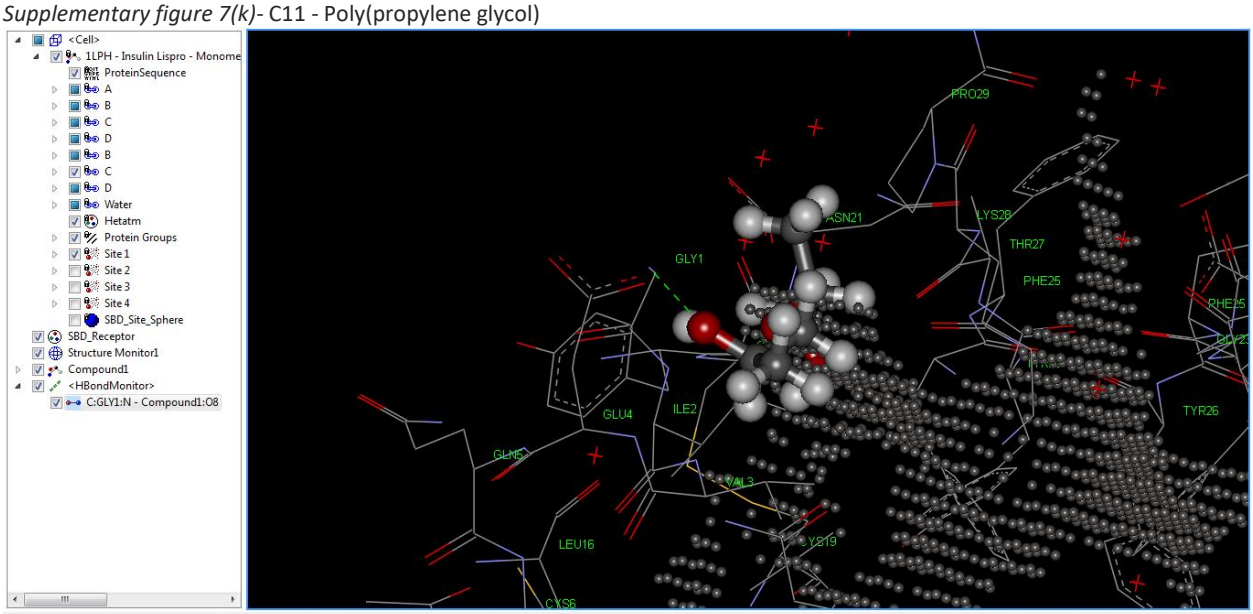
Supplementary figure 7(j)- C10 – Pectin



	Name	ID	Visible	Visibility Locked	Color	Parent	Type	PDB Name	1-Letter	1-Letter & ID#	1-Letter & ID# & Insertion Code	3-Letter	3-Let
52	GLY1	1	Yes	Yes		C	Glycine	GLY	G	G1	G1	Gly	Gly1
53	ILE2	2	Yes	Yes		C	Isoleucine	ILE	I	I2	I2	Ile	Ile2
54	VAL3	3	Yes	Yes		C	Valine	VAL	V	V3	V3	Val	Val3
55	GLU4	4	Yes	Yes		C	GlutamicAcid	GLU	E	E4	E4	Glu	Glu4

Molecule | Atom | AminoAcid | Residue | Bond | Group | BindingSitePoint | BindingSite | ProteinSequence | Cell | AminoAcidChain | Chain | SphereObject | StructureMonitor | HBondMonitor | HBond | Label

Supplementary figure 7(k)- C11 - Poly(propylene glycol)

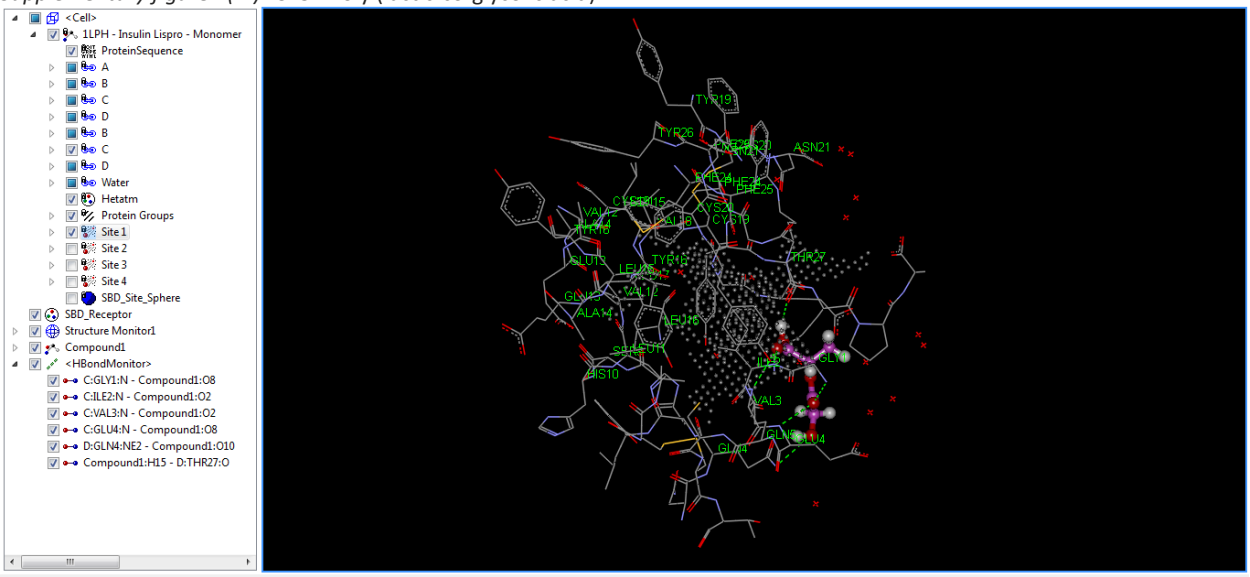


	Name	ID	Visible	Visibility Locked	Color	Parent	Type	PDB Name	1-Letter	1-Letter & ID#	1-Letter & ID# & Insertion Code	3-Letter	3-Let
52	GLY1	1	Yes	Yes		C	Glycine	GLY	G	G1	G1	Gly	Gly1
53	ILE2	2	Yes	Yes		C	Isoleucine	ILE	I	I2	I2	Ile	Ile2
54	VAL3	3	Yes	Yes		C	Valine	VAL	V	V3	V3	Val	Val3

Molecule | Atom | AminoAcid | Residue | Bond | Group | BindingSitePoint | BindingSite | ProteinSequence | Cell | AminoAcidChain | Chain | SphereObject | StructureMonitor | HBondMonitor | HBond | Label

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

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



	Name	ID	Visible	Visibility Locked	Color	Parent	Type	PDB Name	1-Letter	1-Letter & ID#	1-Letter & ID# & Insertion Code	3-Letter	3-Let
99	THR27	27	Yes	Yes		D	Threonine	THR	T	T27	T27	Thr	Thr27
100	LYS28	28	Yes	Yes		D	Lysine	LYS	K	K28	K28	Lys	Lys28
101	PRO29	29	Yes	Yes		D	Proline	PRO	P	P29	P29	Pro	Pro29
102	THR30	30	Yes	Yes		D	Threonine	THR	T	T30	T30	Thr	Thr30

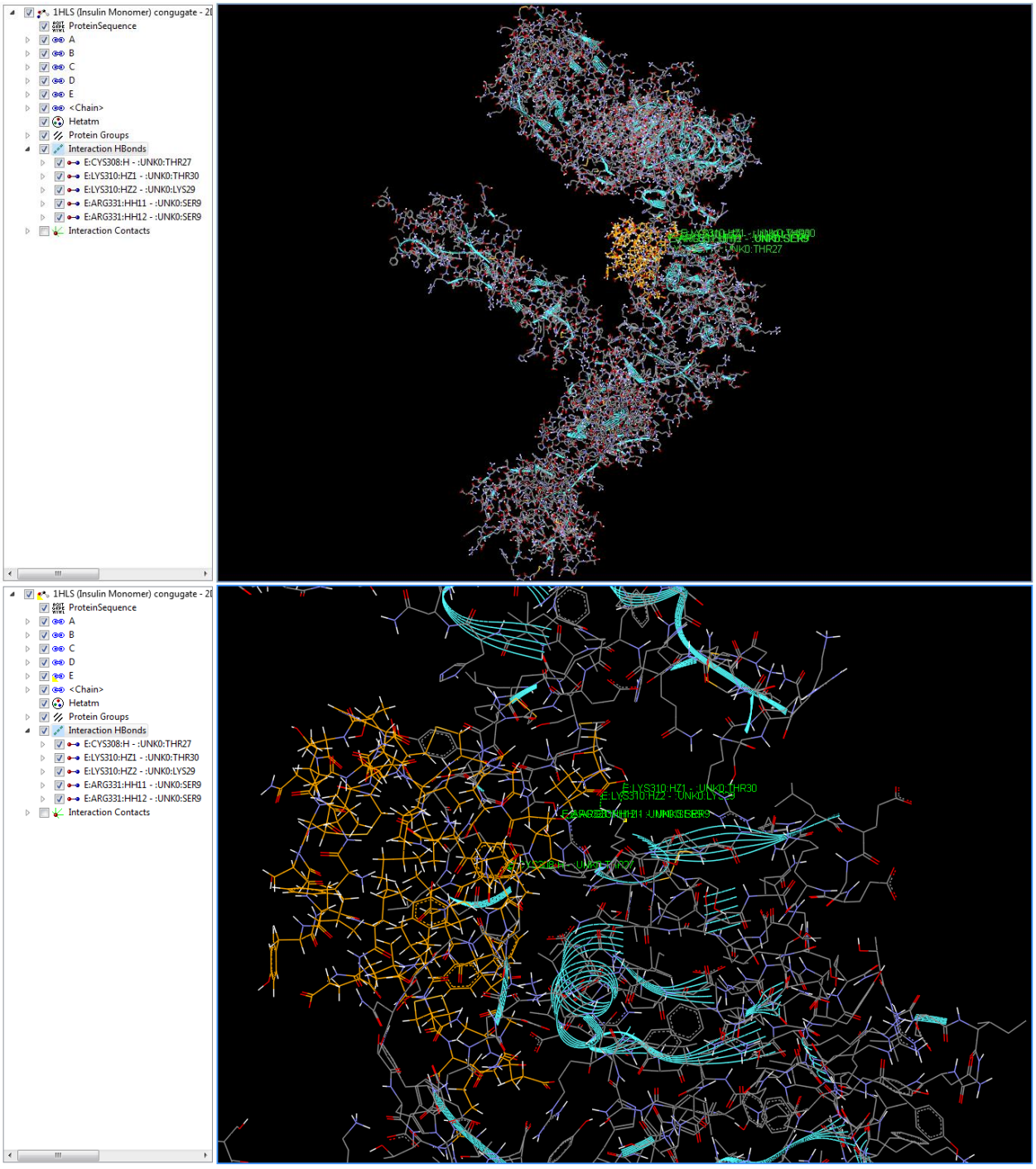
Molecule | Atom | AminoAcid | Residue | Bond | Group | BindingSitePoint | BindingSite | ProteinSequence | Valence | Cell | AminoAcidChain | Chain | SphereObject | StructureMonitor | HBondMonitor | HBond | Label

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

Supplementary figure 8

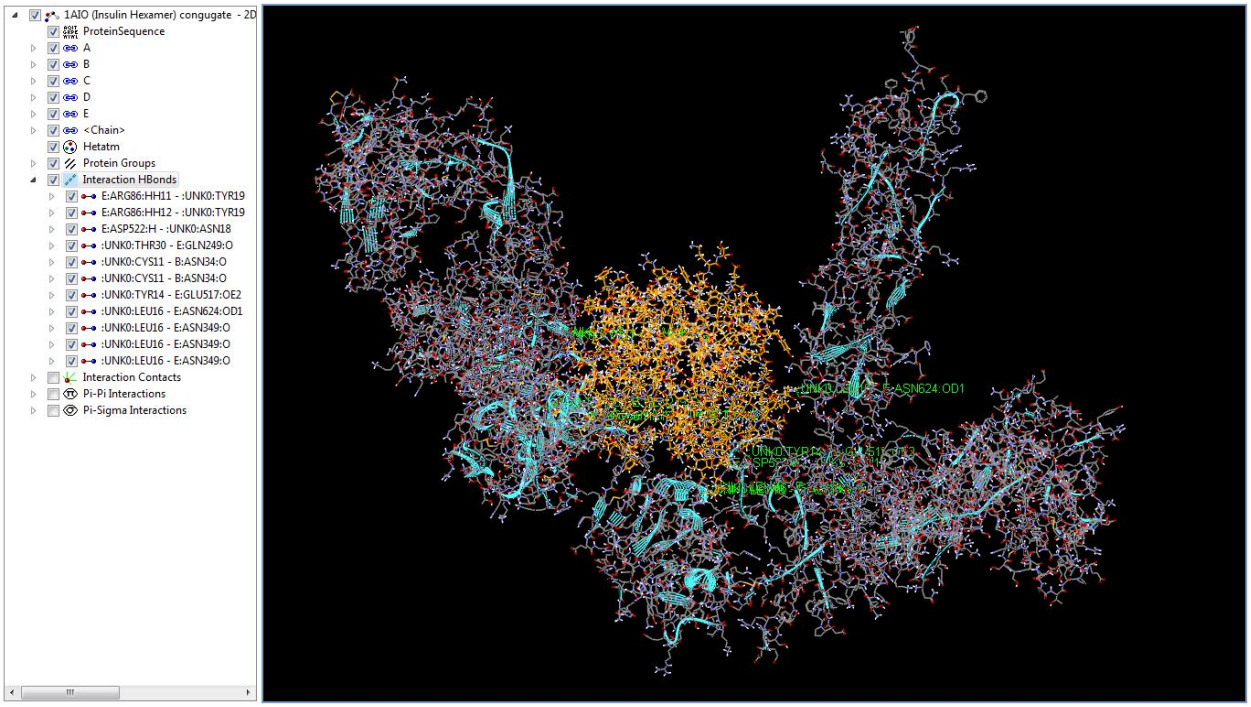
Supplementary figure 8(a)

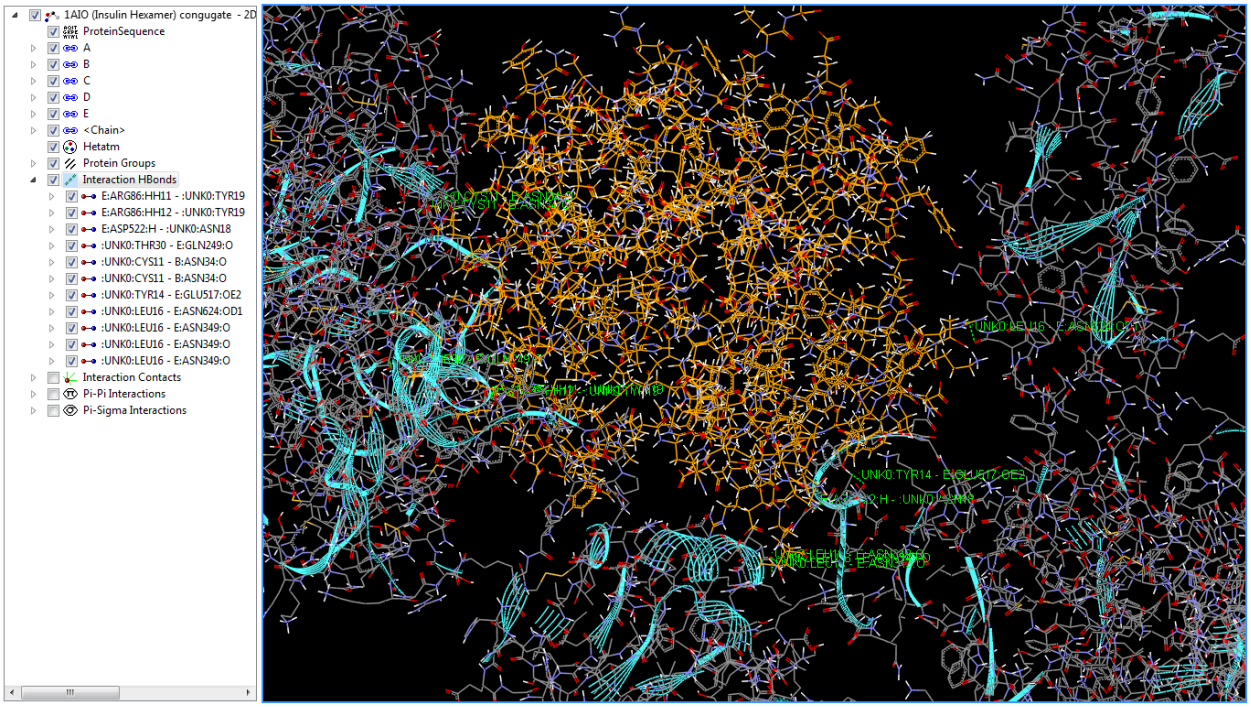
Interaction results of Oral insulin conjugates (Insulin Monomer (1HLS)- DDM Conjugates) with Insulin Receptor (IR). It does not show any interaction in leucine-rich repeat domain (L1, residues 1-157) and in C-terminus of the α-chain (αCT, residues 704-715).



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).