

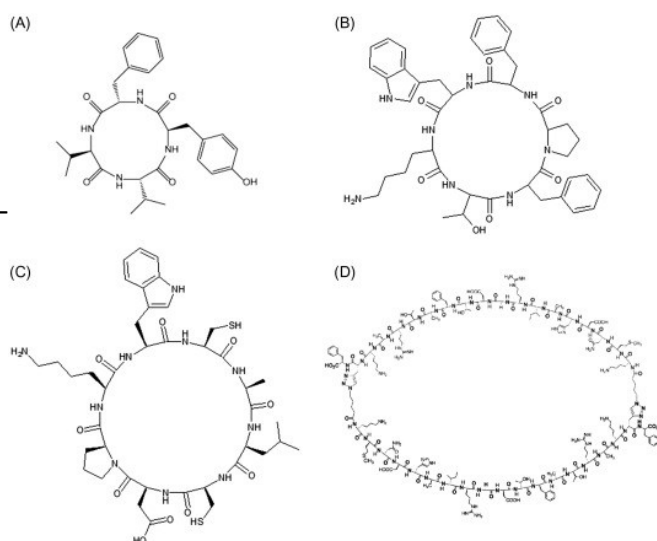
Слайд 2:

В настоящее время актуальной проблемой использования фармацевтических препаратов является то, что все они имеют проблемы проникновения через клеточные мембраны и органеллы или через ГЭБ. Данная проблема может быть решена универсальной системой доставки лекарств: CPP

Слайд 3:

CPP, благодаря своей цикличности, намного сложнее поддаются разложению под действием протеаз в отличие от линейных пептидов. Благодаря этому они и являются перспективными переносчиками лекарств, поскольку сама лекарственная молекула или линейный пептид были бы сразу же атакованы и уничтожены ферментами.

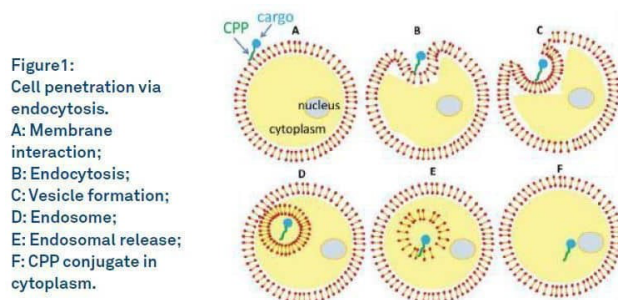
К тому же, они могут присоединять к себе карго – молекулу, которая будет переноситься в мембрану или органеллу клетки за счет CPP



CPP используются в таргетной терапии рака для доставки химиотерапевтических агентов непосредственно к опухолевым клеткам, минимизируя повреждение здоровых тканей.

Они также используются в генной терапии для переноса ген материала в клетки, тем самым позволяя лечить генетические заболевания как, например, муковисцидоз и мышечная дистрофия.

Универсальность и эффективность в преодолении клеточных барьеров делают CPP перспективными в разработке передовых терапевтических технологий и стратегий.



Слайд 4:

Один из первых CPP – TAT. Пептид, выделенный из вируса ВИЧ. Он способен проходить через клеточные мембраны. За счет этой способности, ученые разработали препарат для лечения одного из видов рака: к CPP они 'прикрепили' молекулу, которая, заходя в раковую клетку благодаря проникающему пептиду, разрушает клетку и тем самым убивает рак

Слайд 5:

Конечно, нужно регулировать, в какие клетки препарат будет заходить, а в какие нет, чтобы не атаковать здоровые клетки

Слайд 6:

Проблема вхождения в определенные клетки уже была решена путем линкера, который специфичен только к определенным клеткам

Слайд 8 и 9:

Это наши дальнейшие планы на работу с датасетом. После того, конечно, как мы соберем побольше

Слайд 10:

Без ИИ эту задачу просто не решить, ибо на это требуется огромное количество денег и времени. С моделями МО мы осуществим это намного быстрее

Часть с ИИ:

1) ML:

Методы Машинного Обучения (МО /// Machine Learning = ML), такие как random forests, support vector machines (SVM) и gradient boosting, могут быть использованы для прогнозирования связывания CPP с молекулой мишенью. Эти методы хорошо подходят для обработки числовых и категориальных признаков (например, структура молекулы, заряд, гидрофобность и тд). ML-модели также могут использоваться для классификации молекул: на «подходящие» и «неподходящие» для целенаправленной доставки

Сначала будет проведена подготовка данных. Далее модель МО будет обучена на этих данных, чтобы найти зависимости между характеристиками CPP и молекул, которые обеспечивают успешное связывание и проникновение через мембрану. В итоге модель будет предсказывать, насколько эффективно CPP будет связываться с молекулой-мишенью на основе этих признаков, а также прогнозировать вероятность успешной доставки лекарства в клетку.

Дальше идет описание того, как мы собирали датасет. Я кидал его еще раньше

1) Go to the website on the 1st pic. Sort the peptides by Monomer Length (ascending order). Copy Peptides ID and its SMILES and fill in the sheet with it

2) Go to the web on the 2nd page and past the SMILES in the table, click GO. Look at the molecules in the “Target” and “Compound” columns. If no similar structures found, write NO in the BINDING_DB column

3) Go to the web on the 3d pic and past the peptide SMILES in the “LOAD MOLECULE” table. Choose SMILES in the DESCRIPTOR table. Compare the images of the peptide on the 1st website and this one. If similar, write YES in the “VISUALIZED ON PTB” column

4) From the website on the third step get InChI and past it in the search table on the website from the 2nd step (Binding_DB). If still no similar structures, write No Result in the needed column

5) Always Search by SMILES and InChI on PubChem for cyclic peptides with similar structure. If found, write “found similar str” and past the SMILES/InChI in the needed column on the 2nd list in the google sheet file

6) If on PubChem the peptide with identical or similar structure is found, search on the PubChem for links to Binding_DB

If none found or found but the data there is inappropriate, write N/A in the Binding_Activity column on the 2nd list

Демо-вариант итога работы по ссылке:

<https://docs.google.com/spreadsheets/d/1IV7iuyQE82v3Xyy9LnmvS7FEAKoVbVEwC36WcqJKy0E/edit?usp=sharing>

Изначально ТЗ было такое, но суть работы немного изменилась и в итоге мы искали подходящие данные по немного другим критериям

Конечный вариант на данный момент: тот, который я файлом скинул

Total Conformation Number: refers to the total number of different conformations or spatial arrangements of a molecule that are possible based on its atomic structure and the flexibility of its bonds. It is a measure of the diversity of structural possibilities a molecule can adopt

Unique Conformation Number: It excludes duplicate or equivalent conformations that are the same under symmetry operations or rotations

G_Energy: the lower the energy the more stable the complex

Некоторые из используемых БД:

В файле по ссылке больше инфы о базах



CycPeptMPDB

CyclicPepedia Knowledge Base



Statistics



DataSource



Help ▾



Download

PepBank -- a database of peptides based on sequence text mining and public peptide data sources



The Database of

