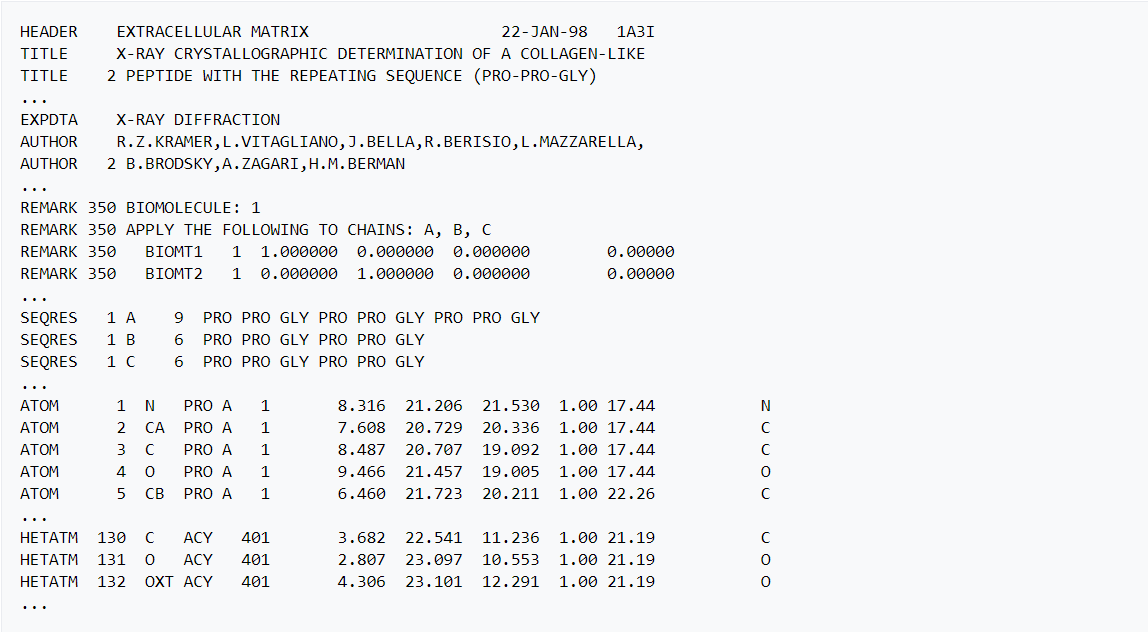
**РОЗДІЛ 2**

**МАТЕРІАЛИ ТА МЕТОДИ ДОСЛІДЖЕННЯ**

**2.1 База даних**

Для аналізу були взяті структури різних білків із білкових банків данних. Структури закодованні у форматі файлу .pdb. Це спеціальний текстовий формат, який описую 3D модель структур закладенних в нього.



HEADER, TITLE та AUTOR – надають інформацію про дослідників, які визначили структуру.

REMARK – містити анотацію у вільній формі, але вони також вміщують стандартизовану інформаціюю.

SEQRES – вказує, які білкові полекули закодовані в цьому файлі, та з скількох і яких амінокислотних залишків складається (3 ланцюга: A, B, C).

ATOM – містить інформацію про всі атоми.

  
На прикладі:

2 – унікальний порядковий індифікатор атома

СА – тип атома

PRO – амінокислотний залишок

А – індифікатор ланцюгу

1 – номер залишку

7.608 20.729 20.336 – координати атома

1.00 – заряд

17.44 – В-фактор

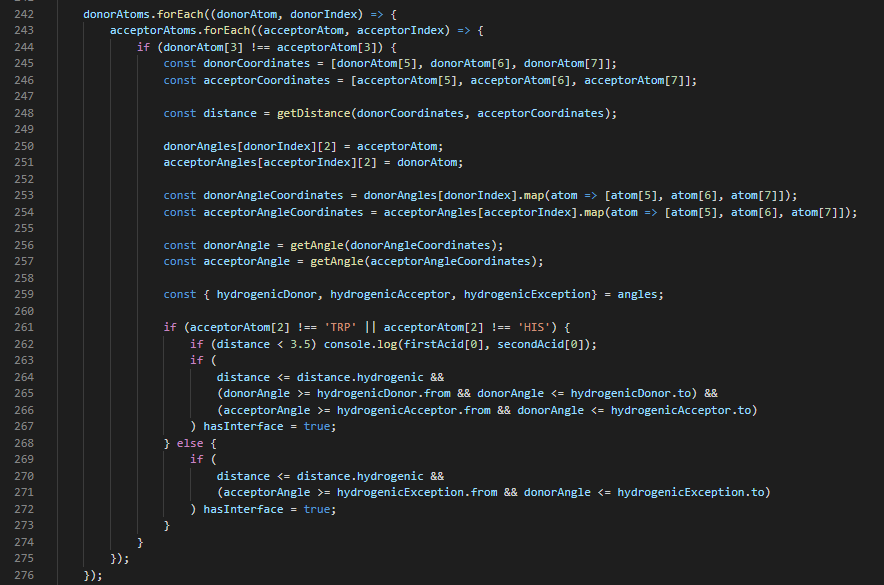
C – узагальнення типу атома

HETATM – опис гетероатомів, тобто тих атомів, які не входять до молекули білка.

**2.2. Мова програмування. Інструменти**

У якості мови програмування був обраний JavaScript (JS). Це високорівнева, динамічна, мультипарадигмальна мова, яка використовується при написанні веб-сайтив та завдяки node.js в серверному та інших загальновживаних умовах.

Node.js – це оточення для виконання JS, побудований на JavaScript-дживку V8. Якраз це оточення і було використано для роботи коду, разом із пакетним менеджером npm, придоставленним node.js, який дозволяє використовувати різноманітні пакети та функція, які були розробленні і викладенні у вільний доступ спілкою JS розробникі для полегшення і пришвидшення роботи.



Приклад невеличкого фрагменту з коду, як виглядає мова програмування JS в IDEA VScode.

Також невідміну частину у розробці відігравали babel поліфіли та компілятори, які також надаються node.js під капот та вручну.

Для контролю версій був використан vcs git. А для моделювання былків PyMol, який побудован на основі язика програмування python.

Та різноманітні модульні пакети, які згадувались вище, наприклад fs для маніпуляцій з файлами, path для полегшення використання абсолютних та відносних шляхів, тощо.

**2.3 Алгоритми та методологія**

У ході розробки даного ПЗ були використанні різноманітні алгоритми, принципи та методології програмування для вирішення низки проблем та задач в коді.

А саме алгоритми швидкого пошуку (quickSearch), швидкого сортування (quickSort), хеш-таблиці, масиви, різноманітні stream методи масивів червоно-чорне дерево, знання про О-велике, Пошук в ширину, алгоритм Дейстри, жадібні алгоритми, динамічне програмування, функціональне програмування, об’єктно-орієнтовне програмування, DRY, KISS та SOLID парадигма, яка вміщує в себе такі ООП принципи: SRP, OCP, LSP, ISP, DIP

**2.4 Методи дослідження**

Методи дослідження базувались на можливості утворення взаємозв’язків між амінокислотами в третинній структурі білку. Так було взято усі амінокислотні залишки одного ланцюгу і перевиренні на можливість утворення різних зв’язків усіма залишками іншого ланцюгу. Перевірялась можливість утворення таких зв’язкі: Pi-pi staking, T staking, Pi-cation staking, Van der Waals, hydrogen bonds.

Pi-pi staking вважалось відбувається між попарно з’єднаними варіаціями ароматичних залишків, а саме PHE, TYR, TRP, HIS. Знаходились координати центроїду кільця молекули, і якщо відстань між двума центроідами була меншою за 4.4А та кут між площинами лежав у межах до 30° - утворюється зв’язок.

T stiking має подібну природу до Pi-pi staking, єдинною різницею являється, що відстань має бути меншою за 5.5А та кут між площинами кілець від 60° до 90°.

Pi-cation staking утворюється між позитивно зарядженим залишком LYS або ARG та ароматичним залишком із групи PHE, TYR, TRP, HIS. Якщо відстань між центроїдом кільця та позитивно зарядженим атомом є меншою за 6.6А та донорський кут від 30° до 90°. На прикладі лізину разяджений атом це NZ, а кут між веторами СЕ-NZ та NZ-центроїд. Для ARG атом - NH1, кут – СZ-NH1 та NH1-цетроїд.

Треба зауважити, що HIS має 2 центроїди. Тому у staking зв’язках перевірялось як можливість утворення зв’язку першим центроїдом так і другим.

Van der Waals порівнювались усі атоми двох гідрофобних залишків, окрім С, СА та СВ. Якщо сумма Ван дер Вальсових радіусів атому та 0.4 була меншою за відстань між атомами, вважалось, що зв’язок утворюється.

Hydrogen bonds утворюється між залишом, який має донорний атом та залишком з акцепторним атомом. Якщо відстань між цими атомами менша 3.5А, донорний кут від 90° до 180°, а акцепторний кут від 80° до 210°.

В якості донорів може виступати: TYR (CZ-OH); SER (CB-OG); THR (CB-OG1); ARG з трьома донорними атомами (CB-NE, CB-NH1, CB-NH2); ASN (CB-NH2); TRP (CE1-NE1); HIS (CE1-NE2); GLN (CD-NE2).

В якості акцепторів: ASN (CB-OD1); TRP (CD2 – NE1); HIS (CG – ND1); GLN (CD-OE1); GLU з двома атомами (CD-OE1, CD-OE2); ASP також з двома (CG-OD1, CG-OD2).

Виключенням є HIS та TRP у ролі донора, їх донорний кут має становити 127°, адже вони міцно закріпленні і тільки в такому випадку можуть утворити водневий зв’язок.

Якщо пара залишків може утворювати PPI зв’язок, який ми шукали, ці залишки відмічались. У разі виявлення більше 10 таких залишків в одному ланцюгу, брався один або більше відповідний фрагмент в залежності від результатів. Цей фрагмент класифікувався за відповідністю до певної вторинної білкової структури. Якщо залишків було менше 10 вважалось, що скоріше за все це вузол або максимум спіраль і така структура не підходить для нашої класифікації.

Фрагменти в залежності від будови можна віднести до alfa-helix, якщо кожен і залишок фрагменту зв’язаний воднево з і + 4 атомом; beta-sheet, якщо кут між векторами залишків CO та ON дорівнює 180° або 135°-140°; beta-turn, якщо якщо кожен і залишок фрагменту зв’язаний воднево з і + 3 атомом.