

**ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ ΜΕ ΕΦΑΡΜΟΓΕΣ ΣΤΗΝ ΒΙΟΙΑΤΡΙΚΗ**

**1η ΕΡΓΑΣΤΗΡΙΑΚΗ ΑΣΚΗΣΗ**

**ΓΙΑ ΤΟ ΕΡΓΑΣΤΗΡΙΟ ΤΟΥ ΜΑΘΗΜΑΤΟΣ:**

**ΒΙΟΠΛΗΡΟΦΟΡΙΚΗ ΙΙ**

**ΕΚΠΟΝΗΘΗΚΕ ΑΠΟ ΤΟΝ ΦΟΙΤΗΤΗ:**

**ΔΙΔΑΣΚΟΥΣΑ:**

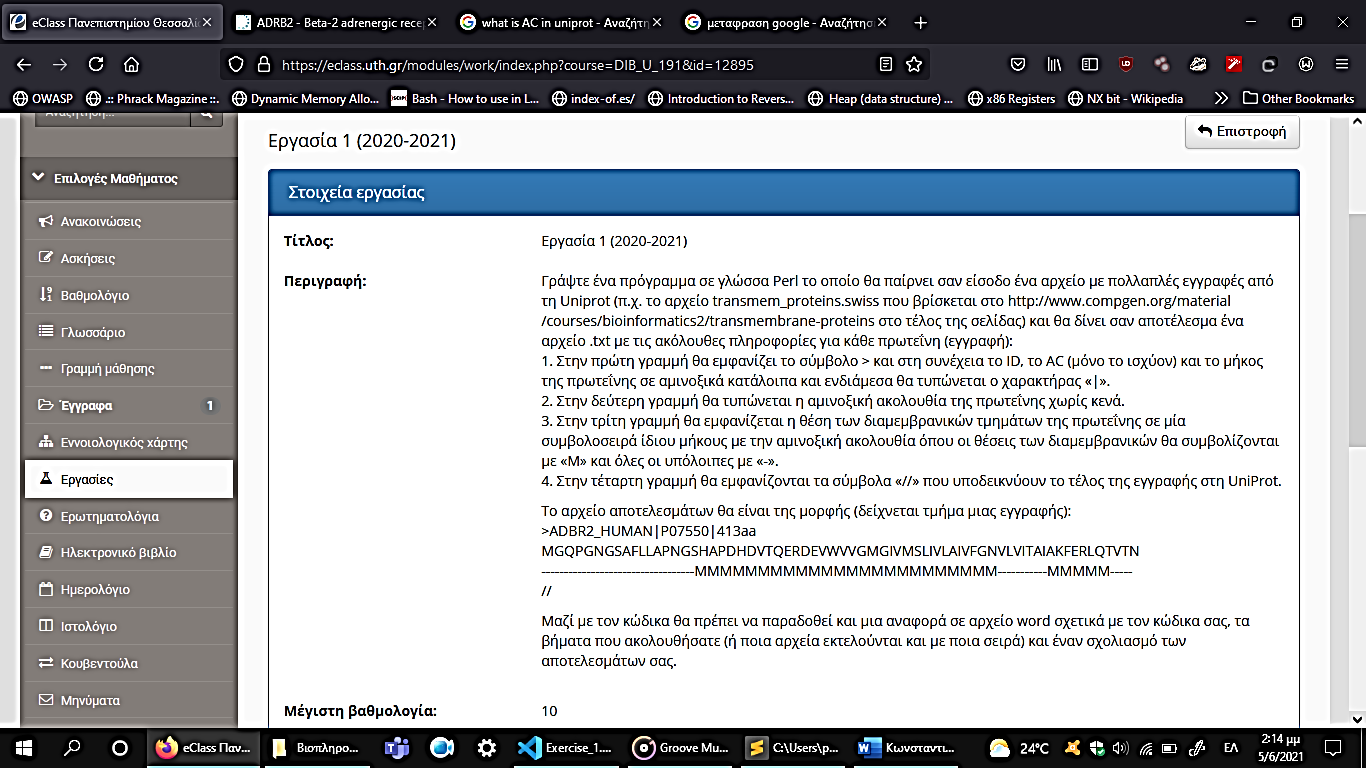
**κ. ΚΟΝΤΟΥ ΠΑΝΑΓΙΩΤΑ**

Λαμία, 2021

**ΠΕΡΙΕΧΟΜΕΝΑ**

* 1. **Εκφώνηση Άσκησης ......................................................................................................σελ. 3**
  2. **Επεξήγηση Κώδικα ...................................................................................................... σελ. 3**
  3. **Παρατηρήσεις ............................................................................................................... σελ. 6**

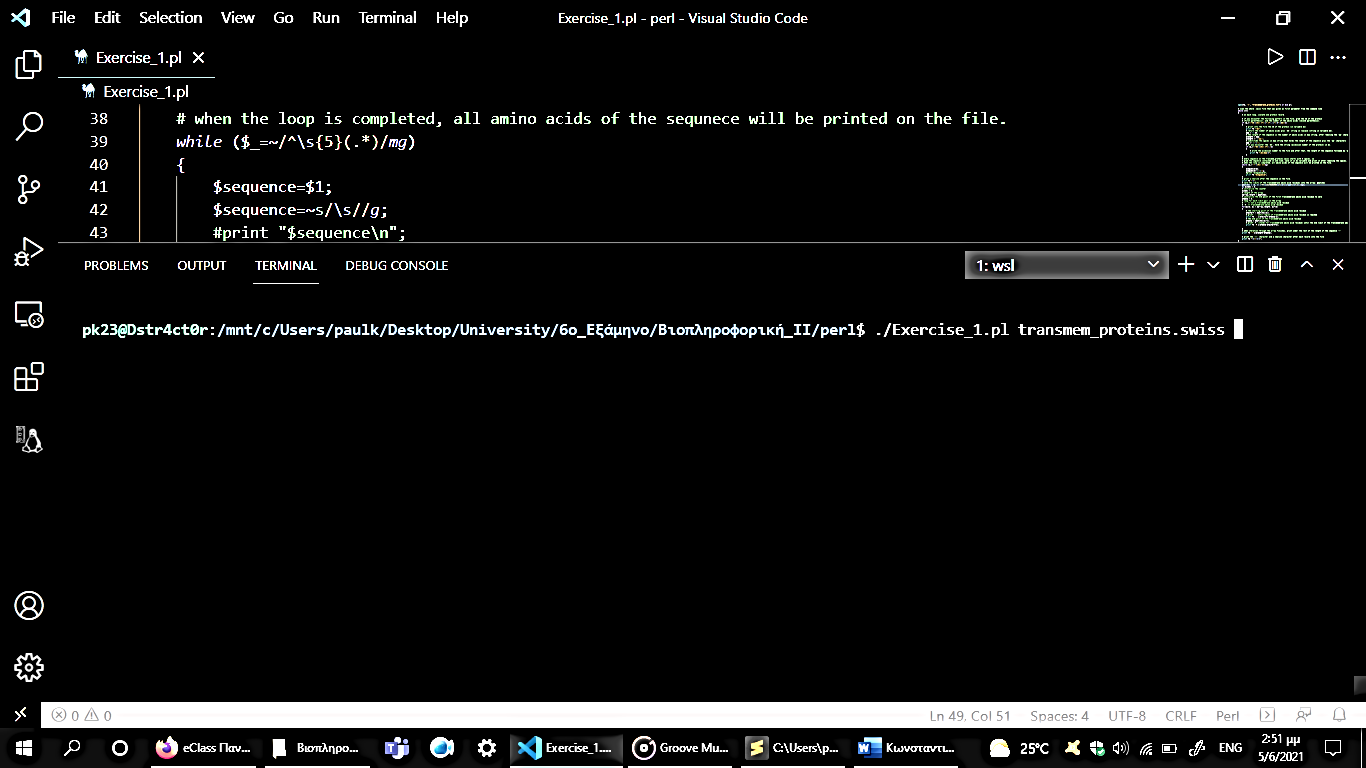
**1.1 Εκφώνηση Άσκησης:**



Εικόνα 1

**1.2 Επεξήγηση Κώδικα:**

Ο κώδικας που υλοποιήθηκε, δέχεται σαν είσοδο ένα αρχείο (απο την γραμμή εντολών ως 1η παράμετρος) .swiss που περιέχει μία ή περισσότερες εγγραφές πρωτεϊνών και στη συνέχεια αποθηκεύει για κάθε εγγραφή, το ID, το AC (Accession Number), το μήκος της ακολουθίας, την ίδια την ακολουθία και τα διαμεμβρανικά κατάλοιπα της ακολουθίας με τη μορφή που φαίνεται παραπάνω (εικόνα 1), στο αρχείο με όνομα «transmembrane\_proteins.txt». Το αρχείο με εγγραφές που χρησιμοποιήθηκε ως παράδειγμα είναι το «transmem\_proteins.swiss». Το πρόγραμμα εκτελείται από την γραμμή εντολών όπως φαίνεται στην παρακάτω εικόνα:

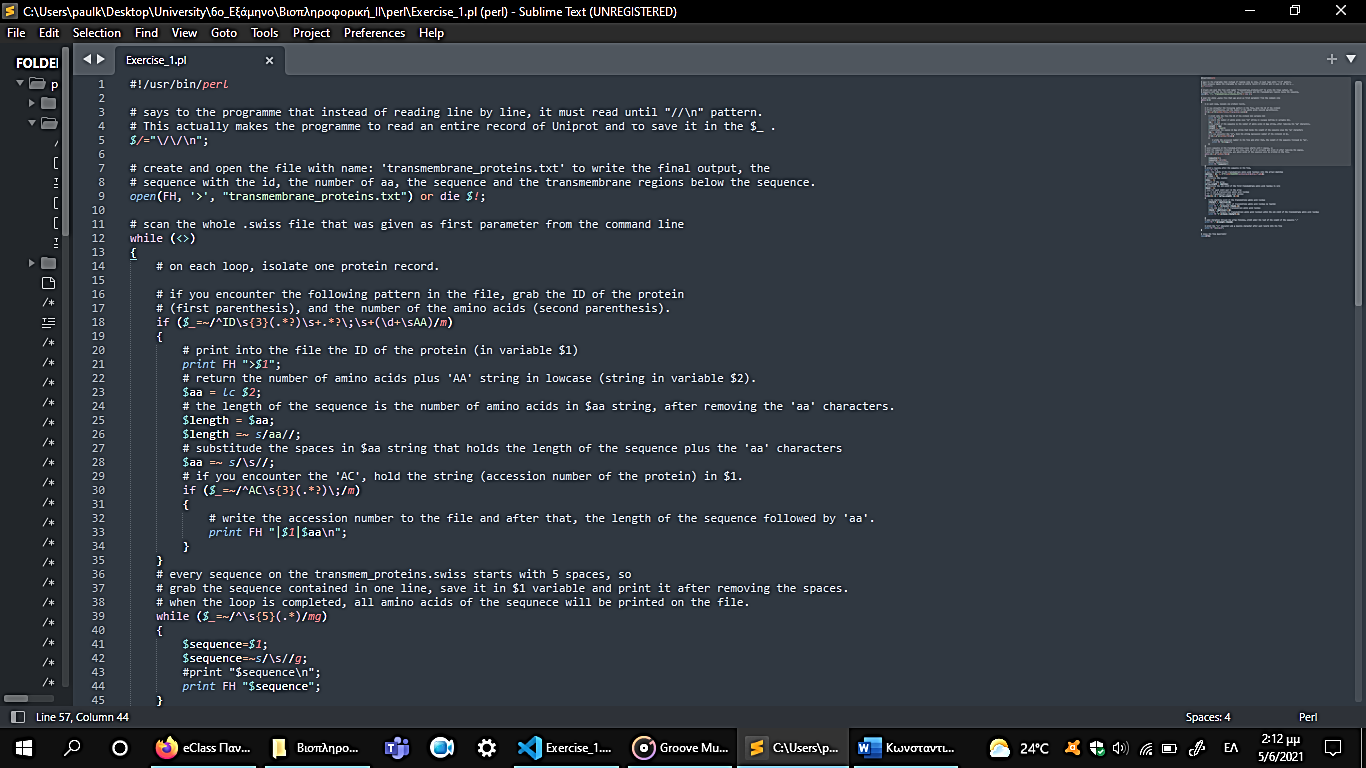


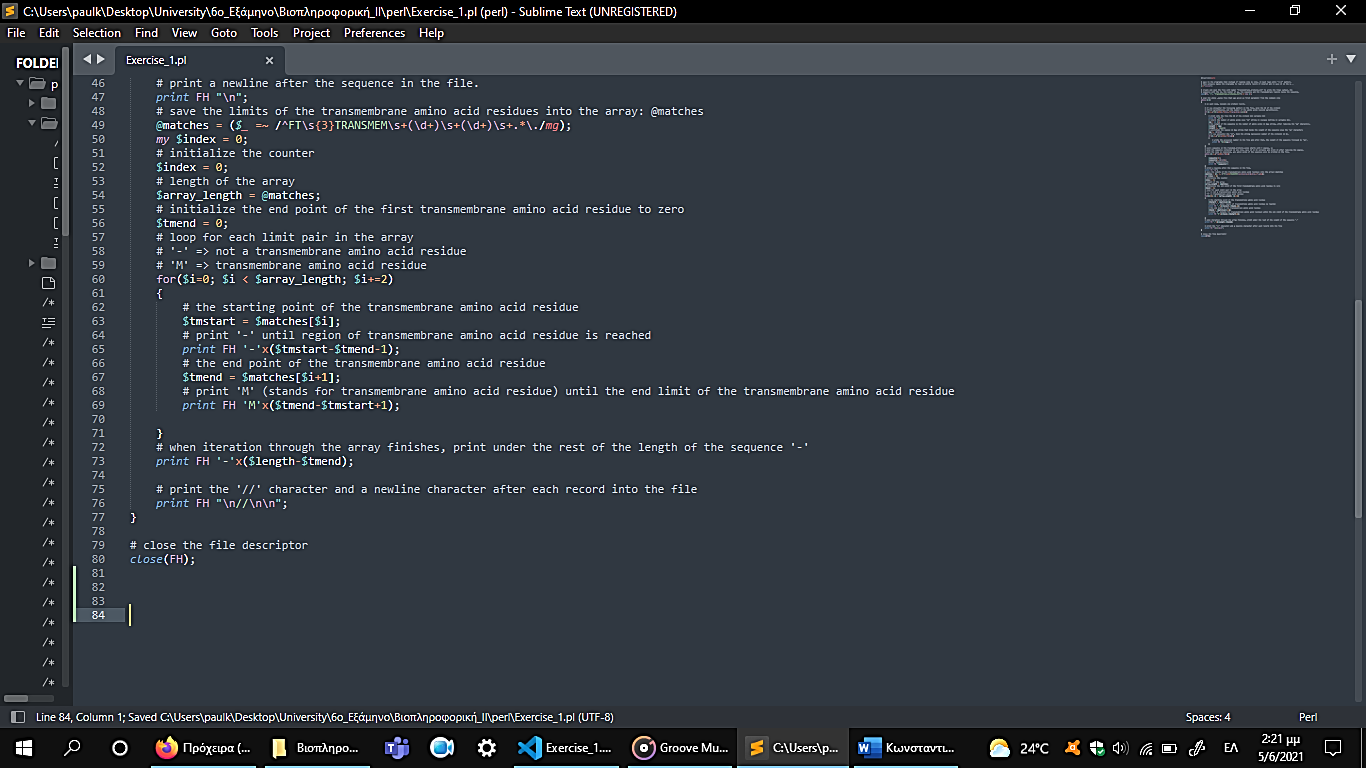
**Αναλυτική Επεξήγηση:**

Αρχικά στη γραμμή 5 σχεδιάζουμε το πρόγραμμα ώστε αντί να διαβάζει γραμμή ανά γραμμή, να διαβάζει μέχρι να εντοπίσει το μοτίβο “//\n”, δηλαδή μέχρι το τέλος της κάθε εγγραφής. Στη συνέχεια, στη γραμμή 9, δημιουργούμε και ανοίγουμε το αρχείο με όνομα «transmembrane\_proteins.txt», για να αποθηκεύσουμε το τελικό αποτέλεσμα δηλαδή τις τροποποιημένες εγγραφές. Με τον επαναληπτικό βρόχο while, «σκανάρουμε» ολο το αρχείο που δόθηκε ως παράμετρος από την γραμμή εντολών. Για κάθε μοτίβο που περιγράφεται στις if (γραμμές 18, 30), επιλέγουμε το ID της κάθε εγγραφής, τον αριθμό προσπέλασης (Accession Number) και το μήκος της πρωτεϊνικής ακολουθίας και τα εκτυπώνουμε στο αρχείο (γραμμές 21,33) διαχωρισμένα με τον χαρακτήρα ‘|’. Έπειτα, στον επαναληπτικό βρόχο while (γραμμές 39-45), απομονώνουμε την ακολουθία που βρίσκεται σε πολλαπλές γραμμές, της αφαιρούμε τα κενά και την εκτυπώνουμε και αυτή στο αρχείο (γραμμή 44).

Εφόσον αποθηκεύσουμε στο αρχείο και την ακολουθία, στην επόμενη γραμμή, θέλουμε να εκτυπώσουμε ‘Μ’ για κάθε αμινοξύ που βρίσκεται εσωτερικά της κυτταρικής μεμβράνης του κυττάρου και ‘-’, για κάθε αμινοξύ που βρίσκεται είτε μέσα στο κύτταρο είτε έξω από αυτο. Για τον λόγο αυτό, αποθηκεύουμε στον πίνακα “@matches” (γραμμή 49), τα όρια των διαμεμβρανικών καταλοίπων. Επιπλέον, αρχικοποιούμε έναν μετρητή (γραμμή 50) και μια μεταβλητή που απθηκεύει το τέλος του κάθε διαμεμβρανικού τμήματος (γραμμή 56), με την τιμή μηδέν. Στη συνέχεια, μέσω ενός επαναληπτικού βρόχου (γραμμές 60 - 71), για κάθε ζεύγος διαδοχικών ορίων διαμεμβρανικών τμημάτων που έχουν αποθηκευτεί στον πίνακα “@matches”, αποθηκεύουμε στην επόμενη γραμμή μετά την αμινοξική ακολουθία στο αρχείο, τον χαρακτήρα ‘-‘ αν το αμινοξύ δεν βρίσκεται εσωτερικά της κυτταρικής μεμβράνης του κυττάρου (γραμμή 65) και τον χαρακτήρα ‘Μ’ διαφορετικά (γραμμή 69). Τέλος, για τα κατάλοιπα που έμειναν, γράφουμε στο αρχείο ‘-‘ όσες φορές χρειάζεται μέχρι το πέρας της ακολουθίας, καθώς γνωρίζουμε ότι δεν βρίσκονται σε κάποιο διαμεμβρανικό τμήμα (γραμμή 73) και εκτυπώνουμε στο αρχείο και τους χαρακτήρες ‘//’ και αυτόν της νέας γραμμής.

Η παραπάνω διαδικασία εκτελείται επαναλαμβανόμενα για κάθε εγγραφή του αρχείου εισόδου, μέχρι το τέλος του. Ο κώδικας παρουσιάζεται παρακάτω:

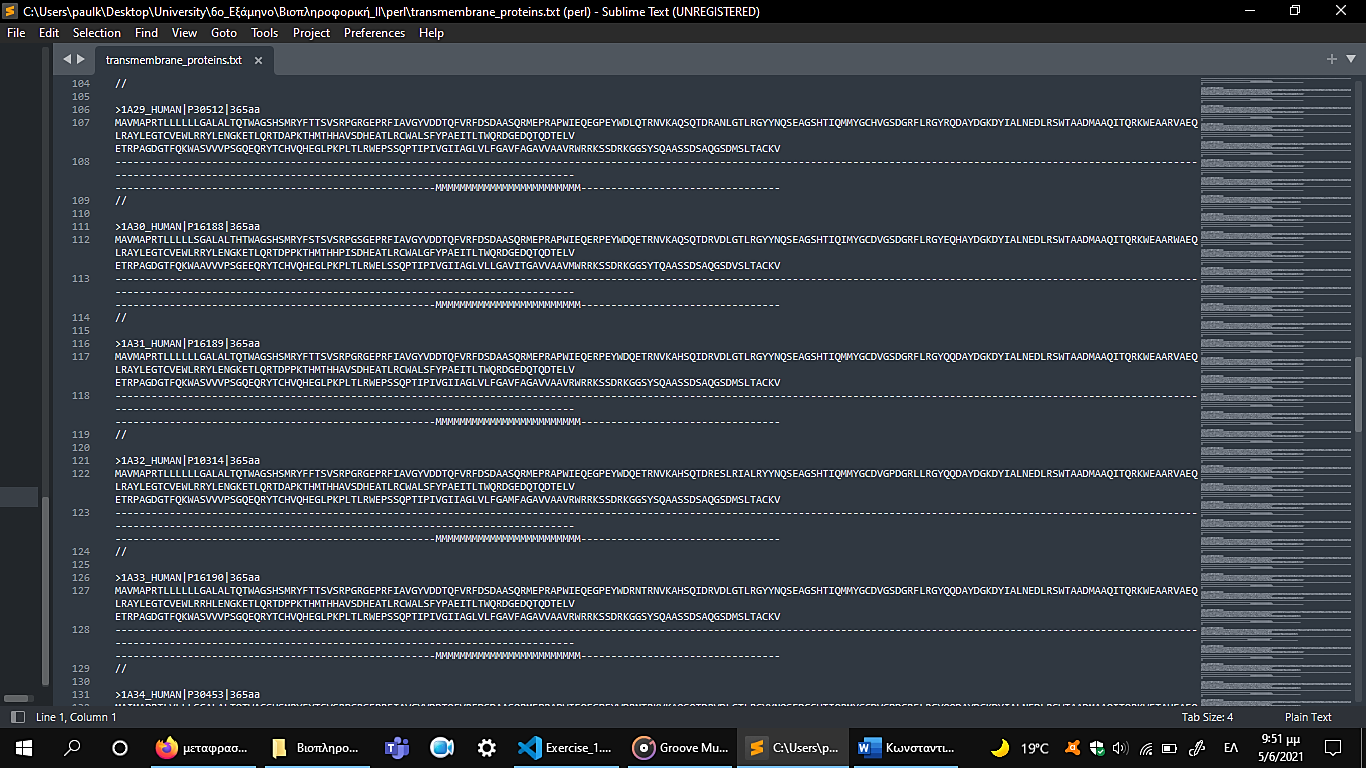




Εικόνες 2 και 3

**1.3 Παρατηρήσεις**

Στην παρακάτω εικόνα (από το αρχείο «transmembrane\_proteins.txt»), παρατηρούμε ότι ορισμένες πρωτεΐνες, έχουν είτε παρόμοιο είτε ίδιο αριθμό διαμεμβρανικών καταλοίπων. Ακόμη, είναι φανερό ότι τα κατάλοιπα αυτά έχουν την ίδια ακολουθία αμινοξέων. Αυτό συμβαίνει διότι οι πρωτεΐνες αυτές προέρχονται από τον ίδιο οργανισμο, στην περίπτωσή μας, τον άνθρωπο.



Εικόνα 4