Σύγκριση δομών πρωτεϊνών

- Η σύγκριση δομών πρωτεϊνών είναι χρήσιμη για την απόδειξη μακρινών εξελικτικών σχέσεων.

- Η δομή αλλάζει σε μικρότερο βαθμό απ’ ό,τι η αλληλουχία των πρωτεϊνών.

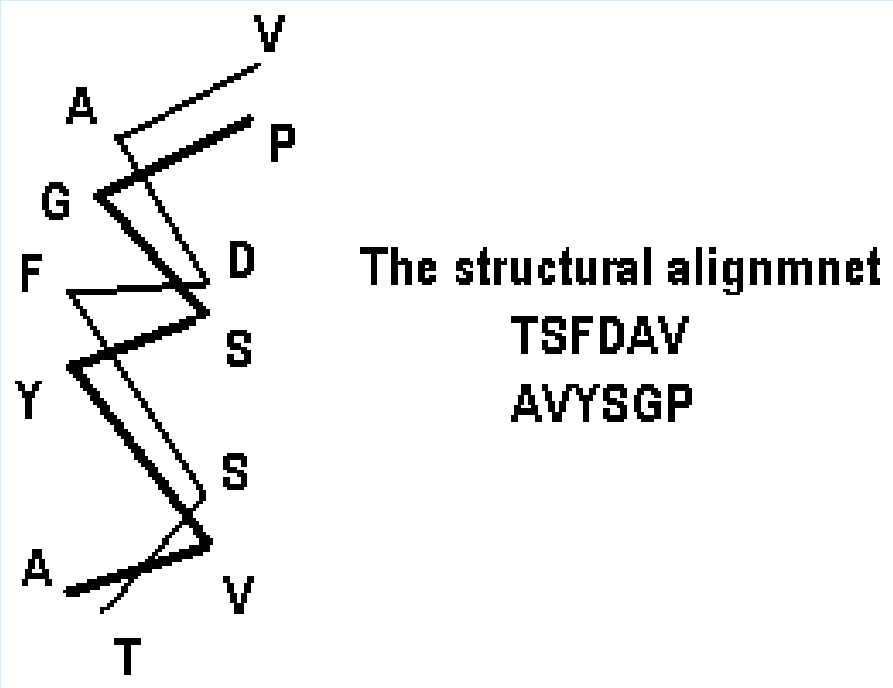
- Μπορεί επίσης να βοηθήσει και στο σχεδιασμό φαρμάκων καθώς βοηθά στην ανάλυση διαμορφωτικών μεταβολών στη θέση με την οποία ο συνθέτης συνδέεται με το υπόστρωμα (conformational changes on ligand binding).

- Δίνει πολύτιμα στοιχεία σχετικά με τις επιπτώσεις που μπορεί να έχει η οποιαδήποτε αλλαγή στις λειτουργίες των πρωτεϊνών.

Αντιστοίχηση δομών

- Σε γενικές γραμμές, η αντιστοίχηση δομών είναι μία διαδικασία κατά την οποία δύο παρόμοιες, τρισδιάστατες δομές τοποθετούνται η μία πάνω στην άλλη, έτσι ώστε οι πεπτιδική αλυσίδα των δομικώς όμοιων αμινοξέα να συμπίπτουν όσο το δυνατόν περισσότερο στο χώρο.

- Αυτή η σύγκριση στη συνέχεια χρησιμοποιείται για τον καθορισμό μίας αντιστοίχησης αλληλουχιών, στην οποία τα αντιστοιχιζόμενα αμινοξέα μοιάζουν δομικά.



Λόγω του ότι κατά την πορεία της εξέλιξης η δομή διατηρείται καλύτερα απ' ό,τι η αλληλουχία, η αντιστοίχηση δομών έχει περισσότερες πιθανότητες να είναι σωστή όσον αφορά τη βιολογική λειτουργία και την εξέλιξη.

- Ακολούθως παρουσιάζονται δύο αντιστοιχήσεις ενός ζεύγους αλληλουχιών σφαιρίνης (globin) με μακρινή σχέση:
  - (a) η πρώτη αντιστοίχηση προέκυψε χρησιμοποιώντας τις αλληλουχίες και
  - (b) η δεύτερη είναι αντιστοίχηση δομών.

- Οι δύο αντιστοιχήσεις, όπως είναι αναμενόμενο, διαφέρουν, με την αντιστοίχηση αλληλουχιών να αποδίδει μεγαλύτερη ομοιότητα ανάμεσά τους.

- Ωστόσο, τα histine (H) αμινοξέα αντιστοιχίστηκαν μόνο στην αντιστοίχηση δομών. Η histine είναι σημαντική, καθώς ελέγχει το σίδηρο (heme iron) και στις δύο σφαιρίνες.

```
 SLSAAEADLAGKSW-APVFANKN-ANGLDFLVALFEKFPDSANFF-ADFKGKSVA--DIK
V-LSPADKTNVKAAWGK-VGAHA-GEYGAEALERMFLSFPTTKTYFPHF-------DLS-H

                                                *
ASPKLRDVSSRIFTRLNEFVNNAANAGKMSA-MLSQ-FAKEHV-GFGVGSAQFENVRSM-F
GSAQVKGHGKKVADALTNAVAHV-D--DMPNAL--SALSDLHAHKLRVDPVNFKLLS-HCL
                                                *

PGFVA
LVTLA
```

Πηγές δεδομένων (Data resources)– PDB

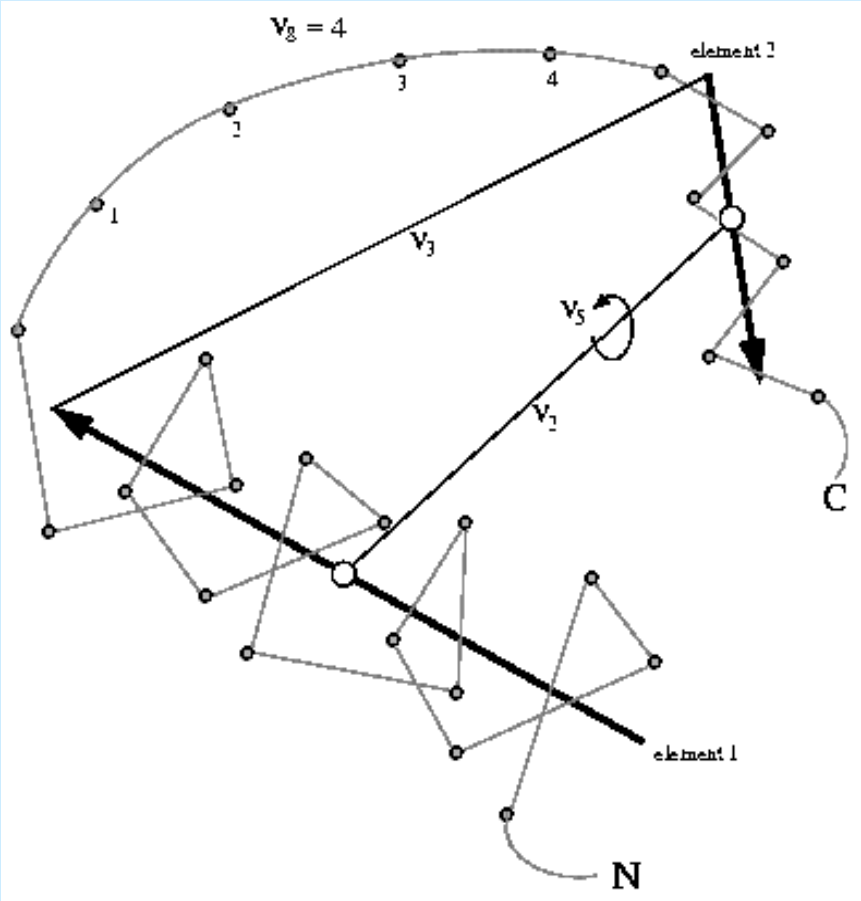
Ο μεγαλύτερος αριθμός δομών πρωτεϊνών βρίσκεται αποθηκευμένος στην Protein Databank (PDB), η οποία συντηρείται στο Research Collaboratory of Structural Biology (RCSB), του Rutgers University και στο EBI.

Αλγόριθμοι σύγκρισης δομών

- Η πλειονότητα των χρησιμοποιούμενων μεθόδων βασίζεται στη σύγκριση των <u>γεωμετρικών ιδιοτήτων</u> και σχέσεων ανάμεσα στα στοιχεία της δευτεροταγούς δομής και των αμινοξέων που βρίσκονται κατά μήκος της ανθρακικής αλυσίδας.

- Οι γεωμετρικές ιδιότητες των αμινοξέα και οι δευτεροταγείς δομές καθορίζονται από τις τρισδιάστατες συντεταγμένες της δομής.

- Στη σύγκριση περιλαμβάνονται και άλλες ιδιότητες, όπως για παράδειγμα <u>φυσικοχημικές ιδιότητες</u> (π.χ. υδροφοβία).



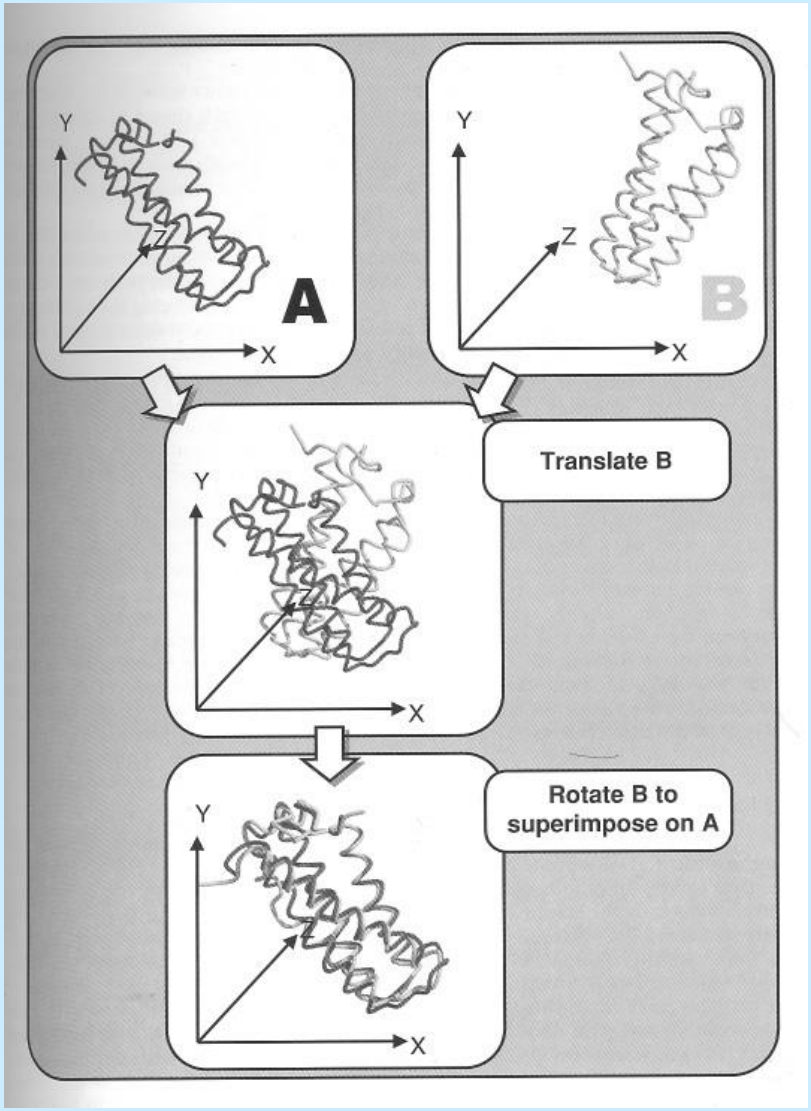
Zintzaras et al. (2006) Comp Biol Med

Μέθοδοι

Οι μέθοδοι για τη σύγκριση δευτεροταγών δομών χωρίζονται σε δύο κατηγορίες: Ενδομοριακές (intermolecular) και διαμοριακές (intramolecular) μέθοδοι.

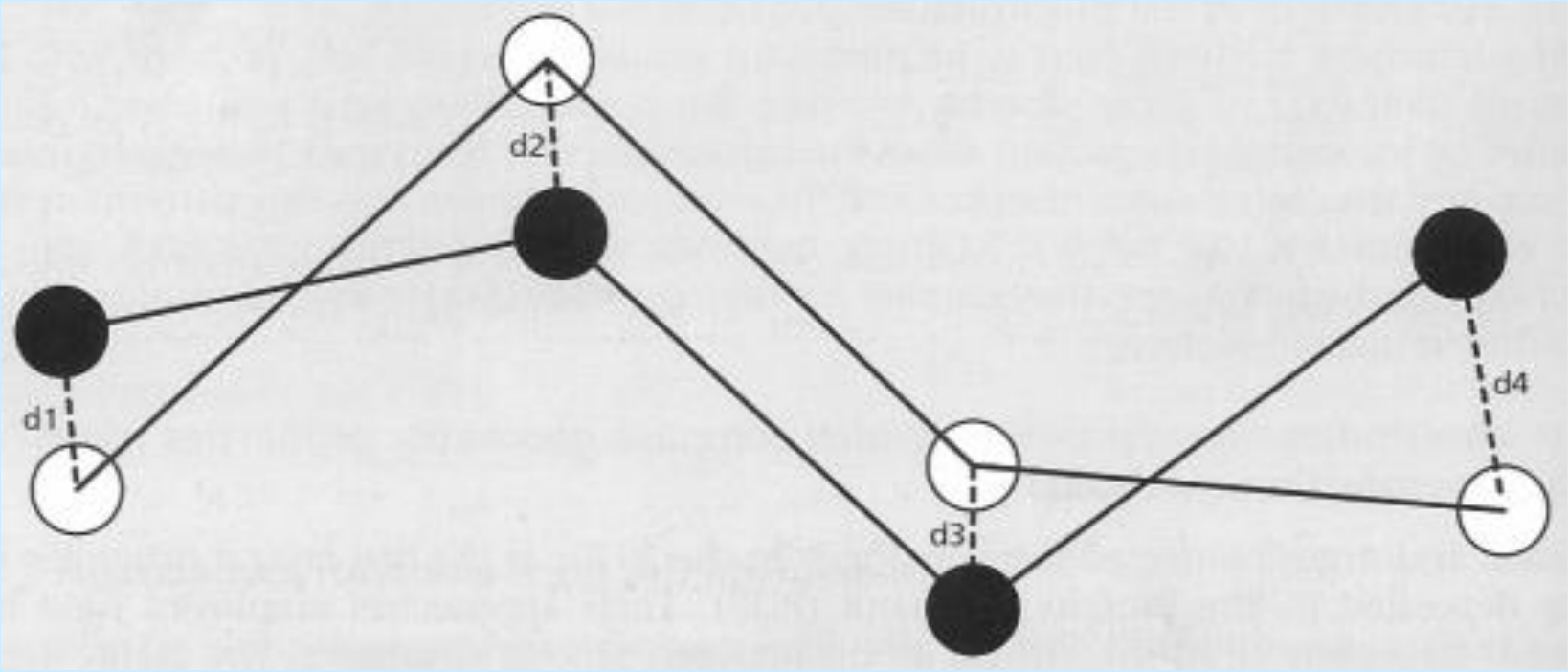
Ενδομοριακές (intermolecular) μέθοδοι

- Οι ενδομοριακές μέθοδοι (rigid body superposition methods) συγκρίνουν, τοποθετώντας τη μία πάνω στην άλλη, δομές πρωτεϊνών και υπολογίζουν τις ενδομοριακές αποστάσεις. Οι μέθοδοι αυτές συγκρίνουν γεωμετρικές ιδιότητες, π.χ. τις θέσεις των αμινοξέα σε τρισδιάστατο χώρο.



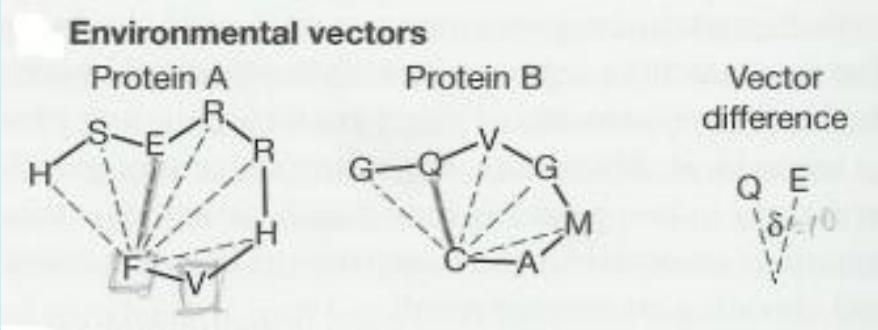
- Η απόσταση ανάμεσα σε δύο αντίστοιχες θέσεις περιγράφεται με τη βοήθεια μίας συνάρτησης διαφορών (residual function) , η οποία υπολογίζει την απόσταση ανάμεσα σε συγκρινόμενα αμινοξέα και η οποία είναι η εξίσωση των ελαχίστων τετραγώνων (Root Mean Squared Deviation – RMSD):

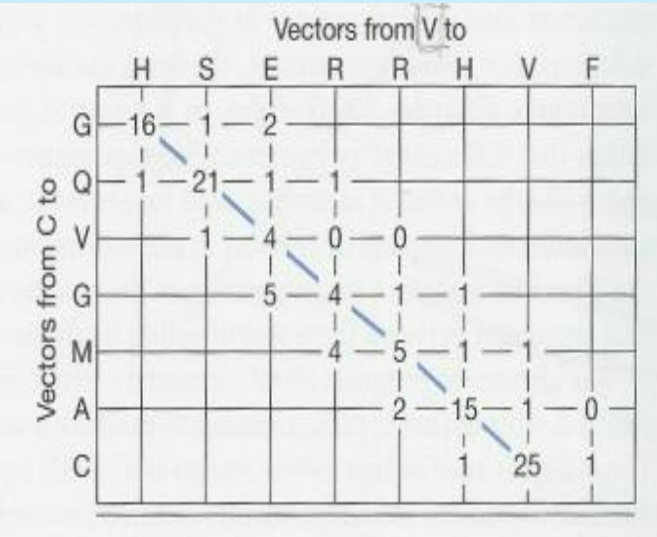
- $RMSD=\sqrt{\Sigma di^2/N}$.

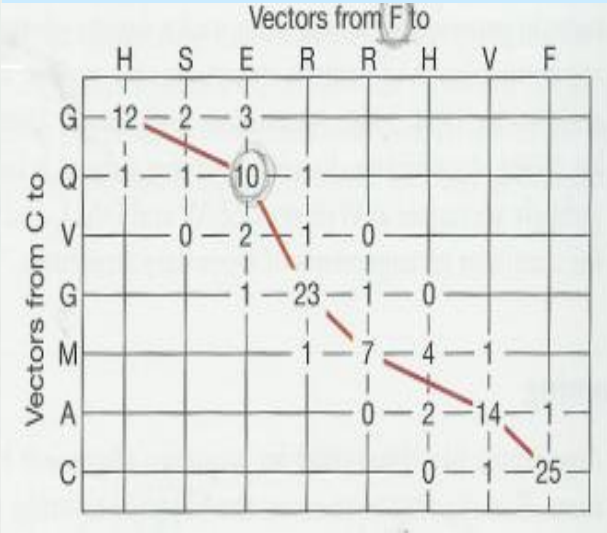


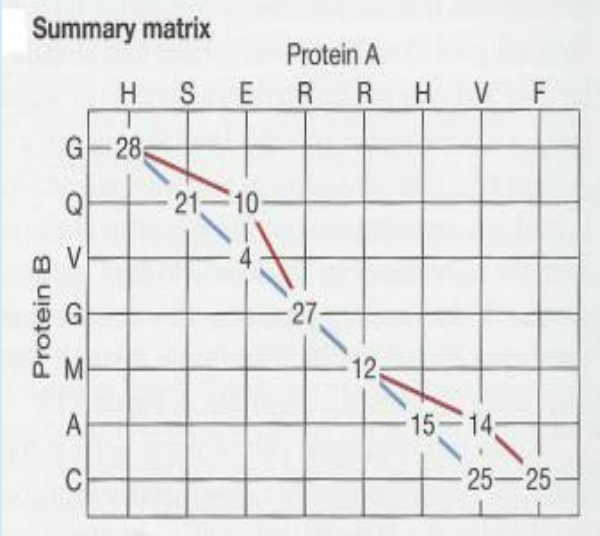
Διαμοριακές μέθοδοι

- Οι διαμοριακές μέθοδοι βασίζονται στη σύγκριση διαμοριακών αποστάσεων ή διανυσμάτων. Αντιστοιχίζουν δομές πρωτεϊνών με βάση πληροφορίες σχετικά με εσωτερικές σχέσεις μέσα σε κάθε πρωτεΐνη.
- Η μέθοδος SSAP (Sequential Structure Alignment Program) είναι μια διαμοριακή μέθοδος για τη σύγκριση δύο δομών πρωτεϊνών, η οποία εφαρμόζει δυναμικό προγραμματισμό σε δύο επίπεδα.









Βάσεις δεδομένων δομών (Structure databases)-CATH

- Η βάση δεδομένων CATH (Class, Architecture, Topology, Homology) περιλαμβάνει μία ιεραρχική κατηγοριοποίηση των δομών των πρωτεϊνών. Η CATH συντηρείται στο UCL. Η κατηγοριοποίηση βασίζεται σε καθολική, μαθηματική και οπτική, σύγκριση δομών.

- **Τα επίπεδα ιεραρχίας της δομικής κατηγοριοποίησης είναι:**

- **Protein class (C):** Αντιπροσωπεύει το ποσοστό των α-ελίκων (α-helices) ή των β-φύλλων (β-strands) μέσα σε μία τρισδιάστατη δομή. Οι βασικές κατηγορίες είναι 3: **mainly-α**, **mainly-β**, **mixed α/β**

- **Protein architecture (A):** Περιγραφή της γενικής κατανομής των δευτεροταγών δομών (α-helices και β-strands) σε τρισδιάστατο χώρο, **χωρίς να λαμβάνονται υπ' όψιν οι δεσμοί** τους. Η καταχώρηση μίας πρωτεΐνη γίνεται με το χέρι και χρησιμοποιώντας απλές περιγραφές για την **κατανομή των δευτεροταγών δομών** (π.χ. barrel, roll, sandwich, κ.λ.π.).

- **Protein topology (T) :** Λαμβάνεται υπ' όψιν ο προσανατολισμός των δευτεροταγών δομών και οι μεταξύ τους δεσμοί. Ομαδοποιούνται χρησιμοποιώντας παραμέτρους οι οποίες έχουν προκύψει εμπειρικά , μέσω αλγορίθμων σύγκρισης δομών. Πρωτεΐνες οι οποίες παρουσιάζουν **δομική ομοιότητα >60%** καταχωρούνται στην ίδια τοπολογία.

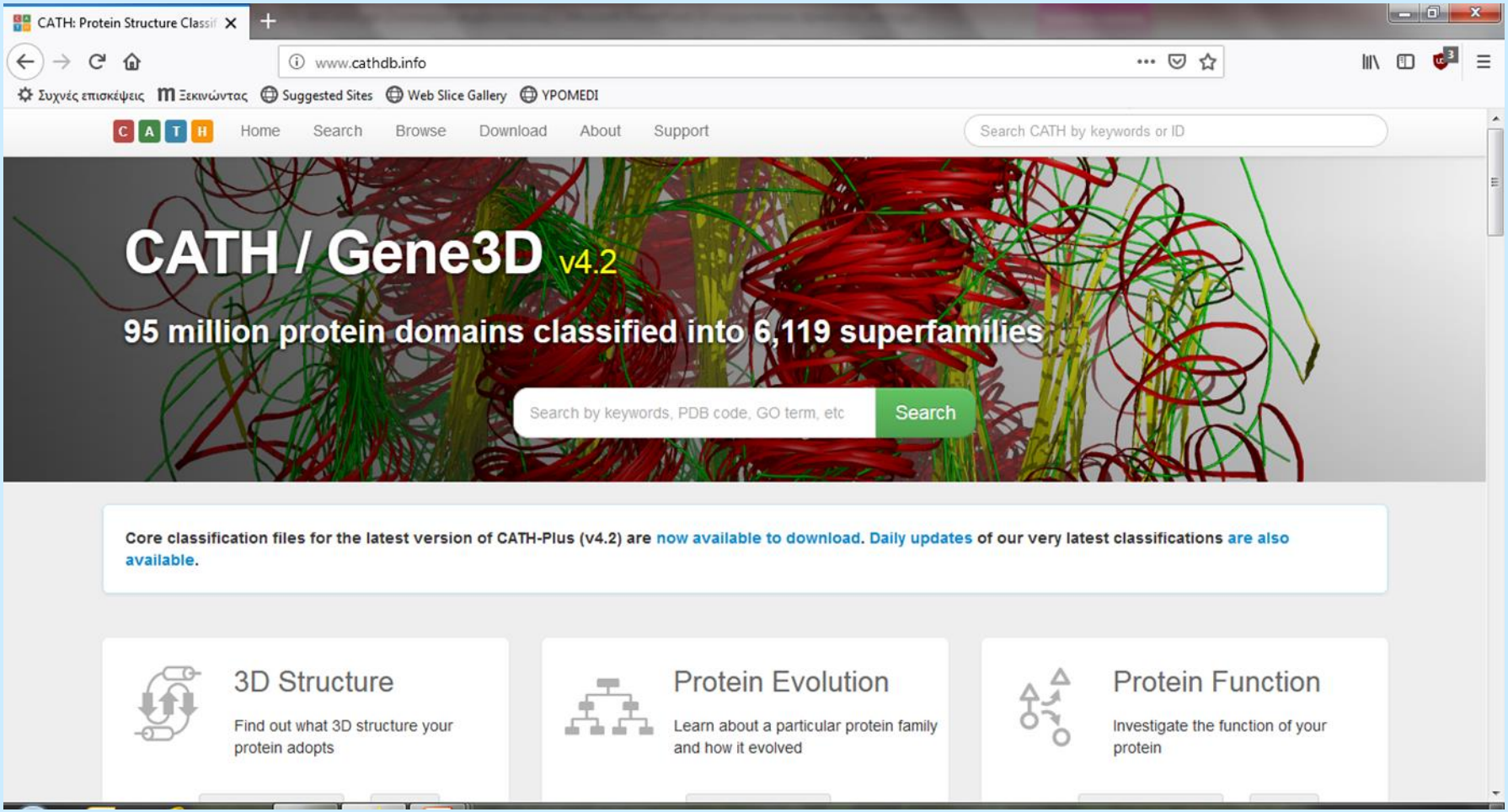
- **Homologous superfamily (H):** Ομάδα πρωτεϊνών των οποίων οι δομές δείχνουν ότι υπάρχει μία κοινή εξελικτική προέλευση (δηλ. ομόλογες). Οι ομοιότητες εντοπίζονται πρωτίστως μέσω της **σύγκρισης αλληλουχιών** και εν συνεχεία μέσω αλγορίθμου **σύγκρισης δομών**.

- **Sequence family:** Δομές μέσα σε ομάδες ομολογίας, μπορούν να διαχωριστούν περαιτέρω με βάση την ομοιότητα των αλληλουχιών τους (ομοιότητες >35%).

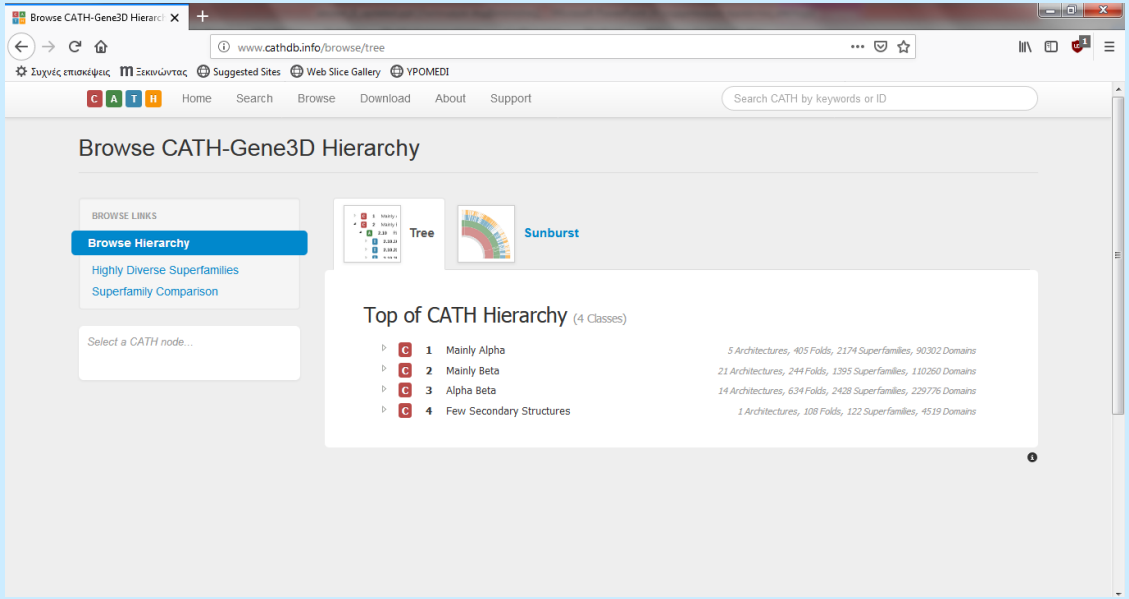
ΠΡΑΚΤΙΚΗ ΕΦΑΡΜΟΓΗ

CATH – Κατηγοριοποίηση δομών (Structure classification)

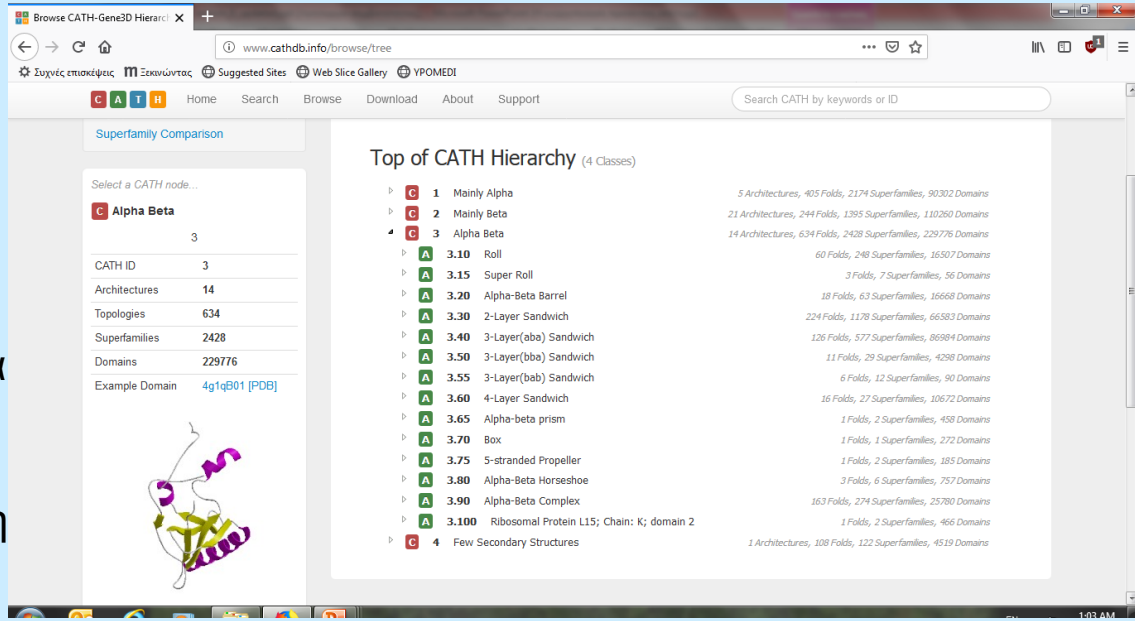
Για να επισκεφθείτε την ιστοσελίδα της CATH, ανοίξτε τον Internet Explorer και πληκτρολογήστε τη διεύθυνση http://www.cathdb.info/. Στη συνέχεια, για να δείτε την ιεραρχική κατηγοριοποίηση συγκεκριμένων πρωτεϊνών, πατήστε στην επιλογή “Browse ”.



Στη νέα σελίδα παρουσιάζονται τα 4 διαφορετικά επίπεδα κατηγοριών (**Classes**). Προκειμένου να δείτε την κατηγορία α+β (μείξη α-β) πατήστε στο “Class 3: Mixed Alpha-Beta”.

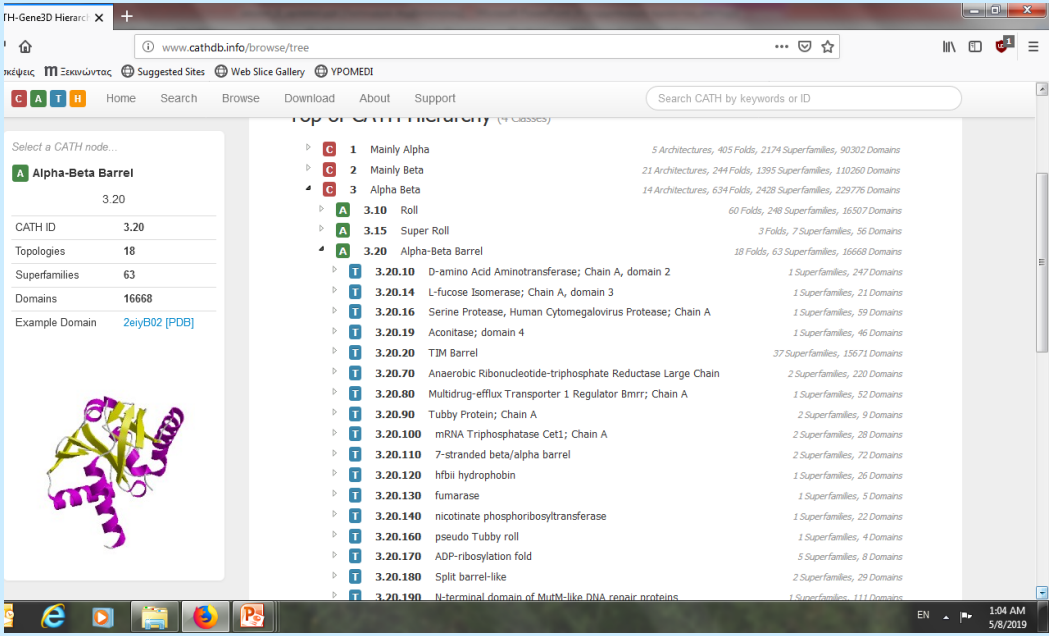


Μπορείτε να δείτε ένα συγκεντρωτικό πίνακα με το πλήθος των εγγραφών για κάθε ένα από τα επίπεδα τα οποία συναντώνται μέσα στην ιεραρχία. Η σειρά των επιπέδων είναι όπως αναφέρθηκε προηγουμένως και το κάθε επίπεδο συμβολίζεται με το αρχικό γράμμα του ονόματός του (A για το Architecture, κ.ο.κ.). Επίσης παρατίθεται ένας πίνακας με τα επίπεδα αρχιτεκτονικής τα οποία συναντώνται μέσα στην τρίτη κατηγορία. Για την τρίτη κατηγορία (class 3) υπάρχουν 14 επίπεδα **Αρχιτεκτονικής (Architecture)**. Πατώντας στο C βλέπετε τα 14 αυτά επίπεδα.

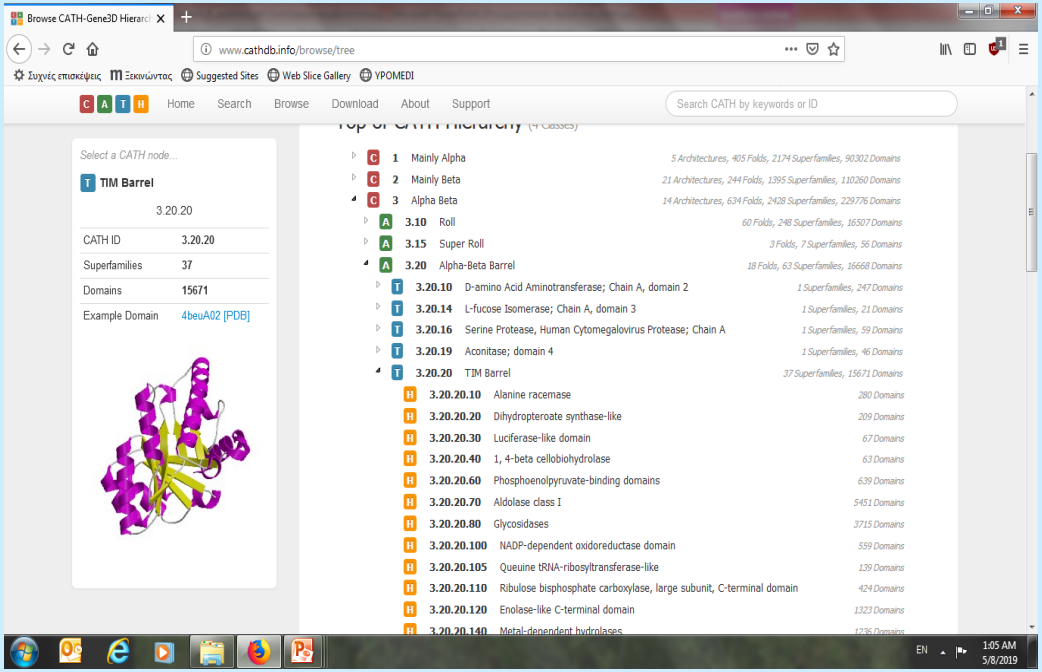


Εάν τώρα επιλέξετε κάποιο επίπεδο και πατήσετε στον αριθμό που αντιστοιχεί στο CATH Code αυτού, θα οδηγηθείτε στην επόμενη κατηγορία που είναι η τοπολογία και όπου παραθέτονται οι πρωτεΐνες οι οποίες ταξινομούνται μαζί με αυτή την οποία επιλέξατε με τα τοπολογικά αυτή τη φορά κριτήρια.

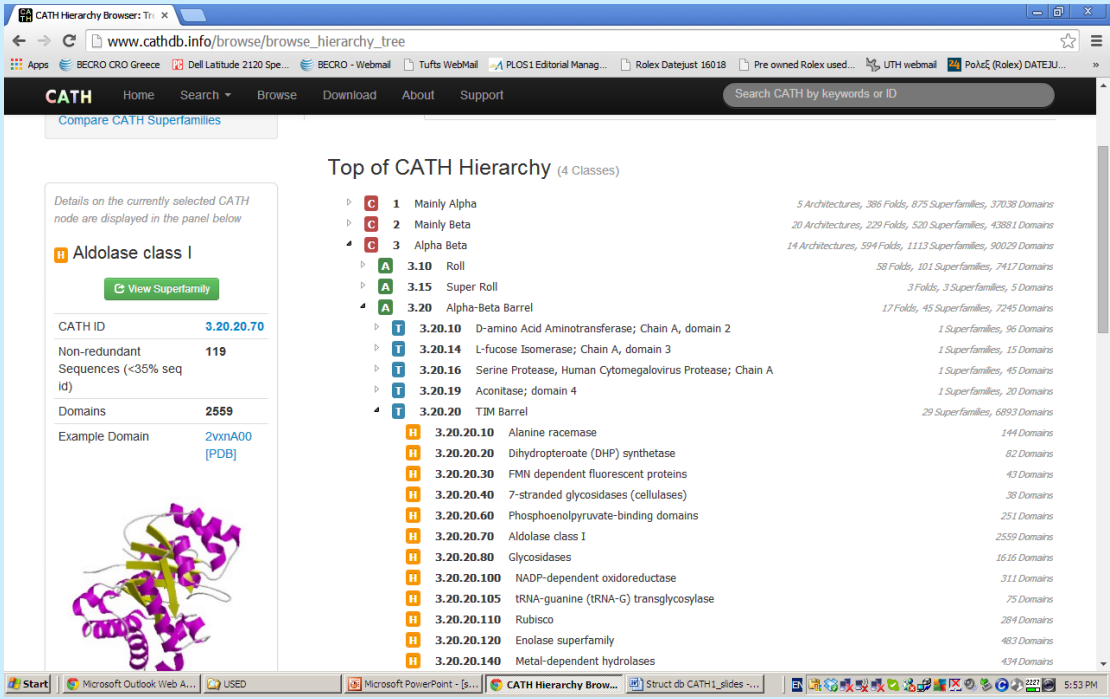
Παραδείγματος χάριν, αν επιλέξετε το επίπεδο ‘Barrel’ και πατήσετε στο 3.20 (CATH Code), τότε θα οδηγηθείτε σε μία σελίδα όπου παρουσιάζονται τα 17 επίπεδα **Τοπολογίας (Topology)** που αντιστοιχούν στην τοπολογία ‘Barrel’.



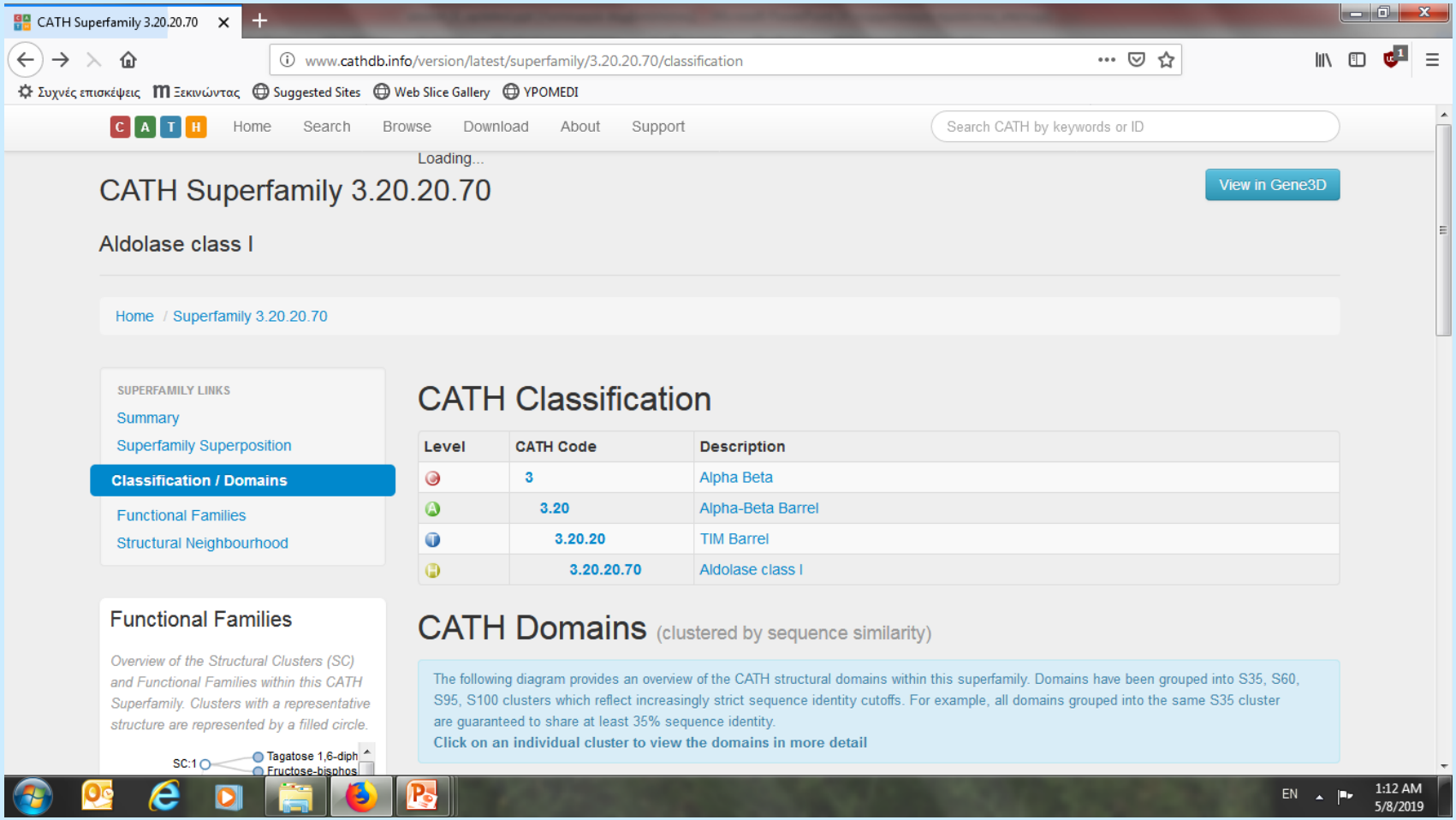
Τώρα, πατήστε στο 3.20.20, TIM Barrel για να δείτε το επόμενο επίπεδο της ιεραρχίας με 29 **Ομόλογες υπερ-οικογένειες (Homologous superfamilies)**.



Πατήστε την 3.20.20.70, Aldolase class I για να δείτε τις 119 αλληλουχίες οι οποίες ανήκουν σε αυτήν την υπερ-οικογένεια ομολογίας.



Οπότε για την Aldolase class I η ταξινομική της είναι η ακόλουθη: .



CATH Superfamily 3.20.20.70

www.cathdb.info/version/latest/superfamily/3.20.20.70/superposition

Συχνές επισκέψεις | Ξεκινώντας | Suggested Sites | Web Slice Gallery | YPOMEDI

C A T H Home Search Browse Download About Support Search CATH by keywords or ID

CATH Superfamily 3.20.20.70

View in Gene3D

Aldolase class I

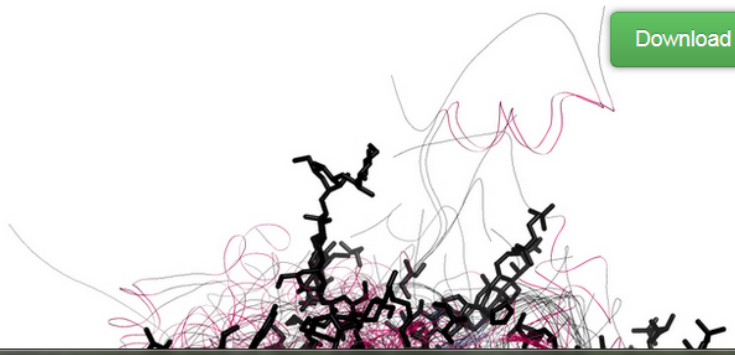
Home / Superfamily 3.20.20.70

SUPERFAMILY LINKS

- Summary
- **Superfamily Superposition**
- Classification / Domains
- Functional Families
- Structural Neighbourhood

Secondary Structure Alignbow

Download as Pymol (49M)

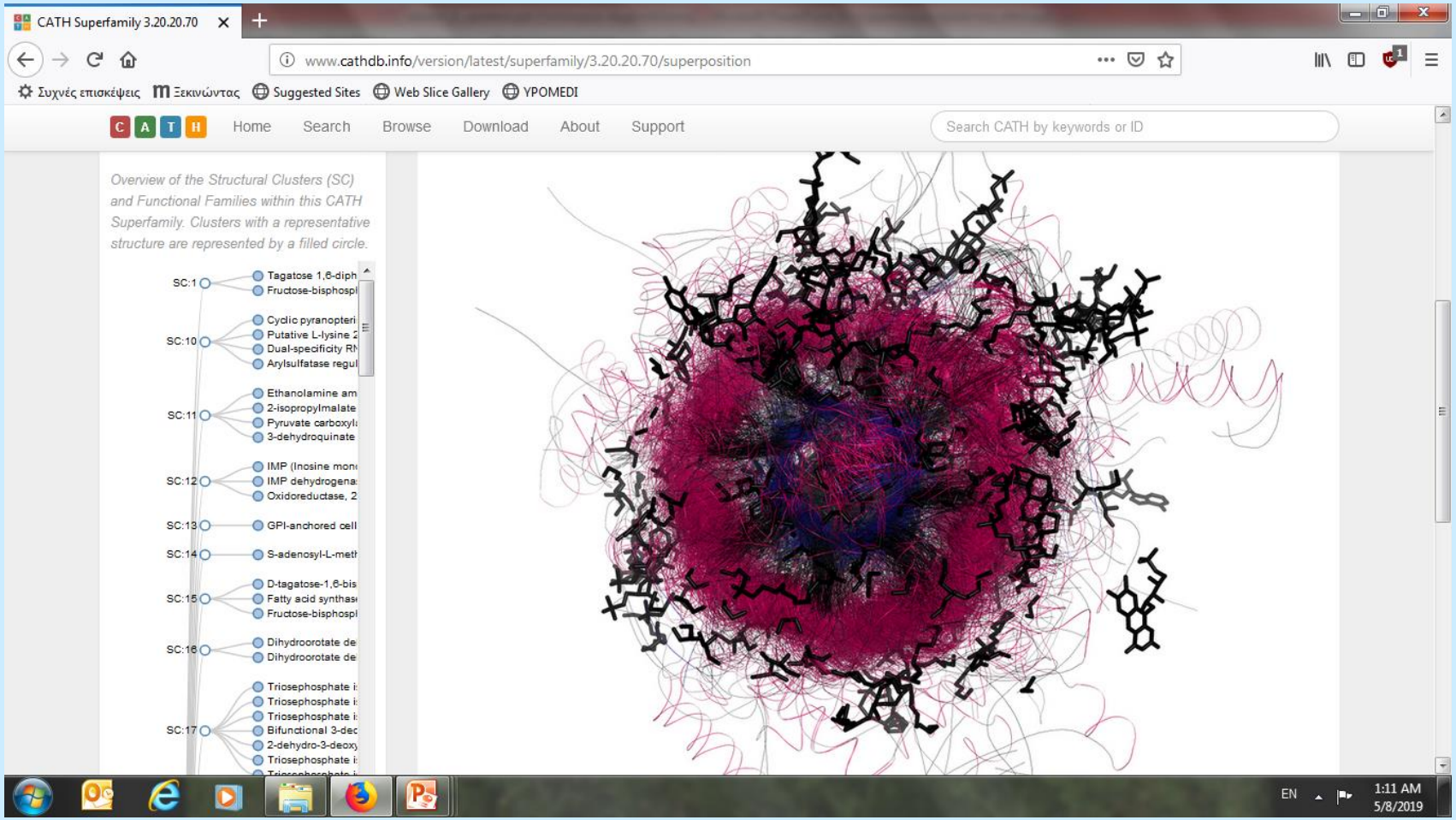


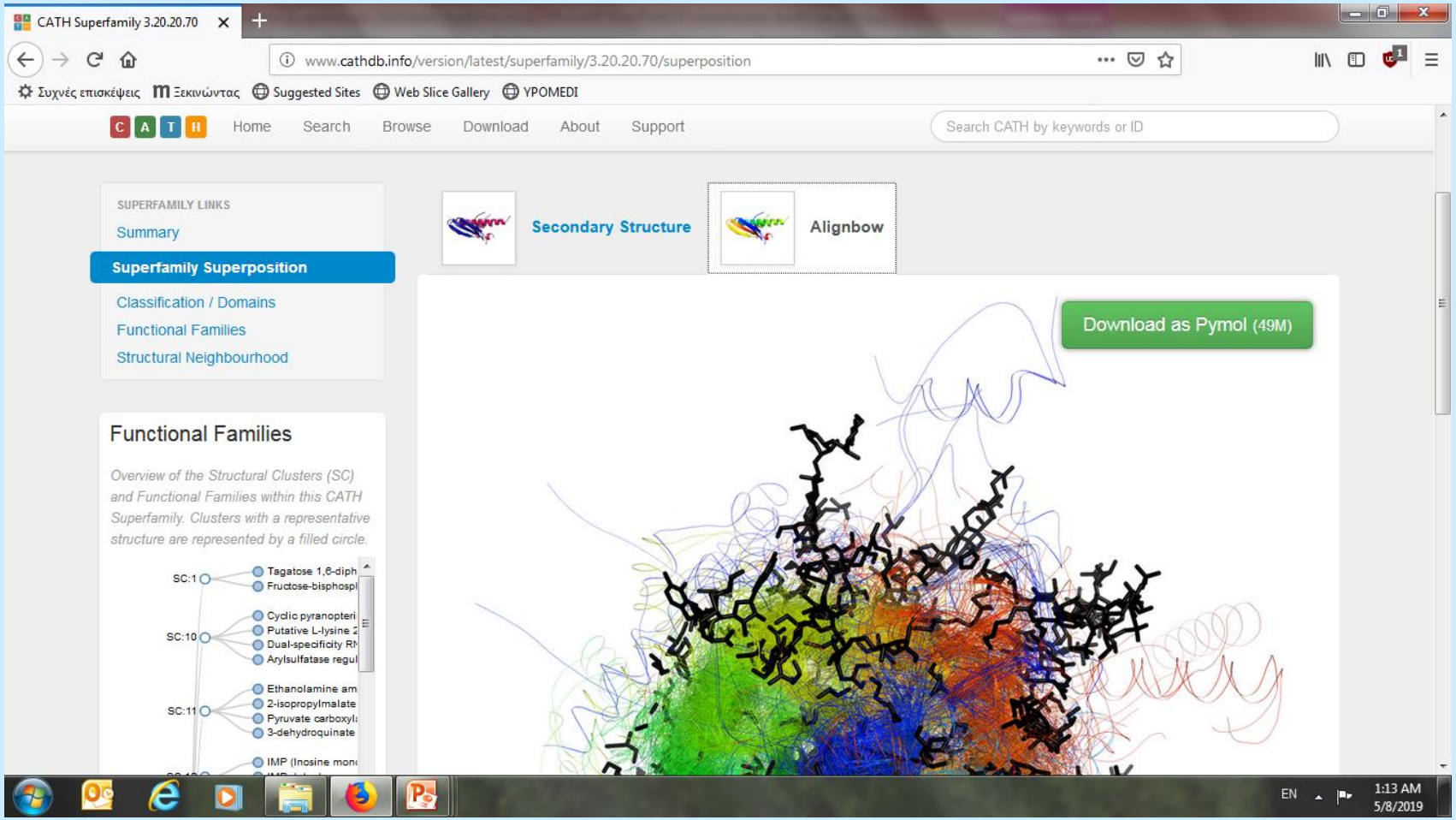
Functional Families

Overview of the Structural Clusters (SC) and Functional Families within this CATH Superfamily. Clusters with a representative structure are represented by a filled circle.

SC:1 — Tagatose 1,6-diph… / Fructose-bisphosp…

EN 1:12 AM 5/8/2019


CATH Superfamily 3.20.20.70
www.cathdb.info/version/latest/superfamily/3.20.20.70/superposition
Suggested Sites
Web Slice Gallery
YPOMEDI
Home
Search
Browse
Download
About
Support
Search CATH by keywords or ID
Overview of the Structural Clusters (SC) and Functional Families within this CATH Superfamily. Clusters with a representative structure are represented by a filled circle.
SC:1
SC:10
SC:11
SC:12
SC:13
SC:14
SC:15
SC:16
SC:17
EN
1:11 AM
5/8/2019


CATH Superfamily 3.20.20.70
www.cathdb.info/version/latest/superfamily/3.20.20.70/superposition
Suggested Sites
Web Slice Gallery
YPOMEDI
Home
Search
Browse
Download
About
Support
Search CATH by keywords or ID
SUPERFAMILY LINKS
Summary
Superfamily Superposition
Classification / Domains
Functional Families
Structural Neighbourhood
Secondary Structure
Alignbow
Download as Pymol (49M)
Functional Families
Overview of the Structural Clusters (SC) and Functional Families within this CATH Superfamily. Clusters with a representative structure are represented by a filled circle.
SC:1
SC:10
SC:11
EN
1:13 AM
5/8/2019

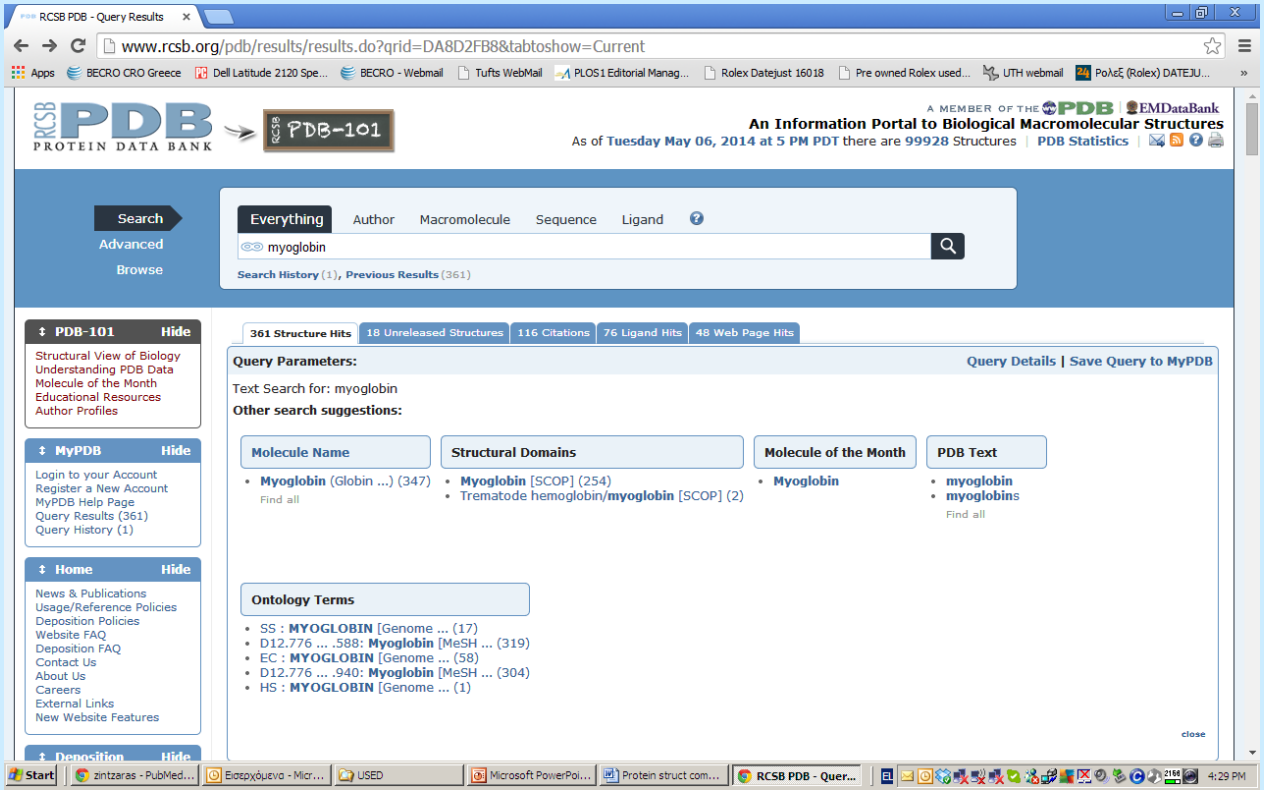
ΠΡΑΚΤΙΚΗ ΕΦΑΡΜΟΓΗ

Σύγκριση δομών πρωτεϊνών – Μέθοδος SSAP

Προκειμένου να συγκρίνουμε τις δομές δύο πρωτεϊνών, π.χ. δύο myoglobins, αρχικά πρέπει να προσδιορίσουμε τους κωδικούς που τους αντιστοιχούν στην Protein Data Bank (PDB).

Για να επισκεφθείτε την ιστοσελίδα της PDB, ανοίξτε τον Internet Explorer και πληκτρολογήστε τη διεύθυνση www.rcsb.org/pdb/ . Στη συνέχεια πληκτρολογήστε στο κενό πεδίο τη λέξη ‘myoglobin’ και πατήστε ‘Search’.

Αυτό θα σας οδηγήσει σε μία νέα σελίδα όπου παραθέτονται όλες οι myoglobins συνοδευόμενες από τους κωδικούς τους και όλες τις πληροφορίες σχετικά με τις πρωτεΐνες.



Πατώντας στον κωδικό πρωτεΐνης ‘1MBN’, εμφανίζονται περιληπτικές πληροφορίες σχετικά με αυτήν την πρωτεΐνη.

RCSB PDB - Query Results

www.rcsb.org/pdb/results/results.do?qrid=1A0635A6&tabtoshow=Current

Showing 1 - 25 of 361 Results

Results : 25 Page: 1 of 15

Filter: Check All View: Detailed Reports: Select one... Sort: Relevance

1MBN The stereochemistry of the protein myoglobin

Authors: Watson, H.C., Kendrew, J.C.

Release: 1976-05-19

Experiment: X-RAY DIFFRACTION with resolution of 2.00 Å Residue Count 153

Compound: 1 Polymer [Display Full Polymer Details | Display for All Results]
2 Ligands [Display Full Ligand Details | Display for All Results]

Citation: PubMed ID is not available.

Molecule of the Month: Myoglobin, PDB Pioneers

Search Hit: Classification: OXYGEN STORAGE (Molecule of the Month:**Myoglobin**, PDB Pioneers)

1MCY SPERM WHALE MYOGLOBIN (MUTANT WITH INITIATOR MET AND WITH HIS 64 REPLACED BY GLN, LEU 29 REPLACED BY PHE

Authors: Li, T., Phillips Jr., G.N.

Release: 1995-12-07

Experiment: X-RAY DIFFRACTION with resolution of 1.70 Å Residue Count 154

Compound: 1 Polymer [Display Full Polymer Details | Display for All Results]
3 Ligands [Display Full Ligand Details | Display for All Results]

Citation: **A double mutant of sperm whale myoglobin mimics the structure and function of elephant myoglobin.** (1995) J.Biol.Chem. **270:** 20763-20774 [Display Full Abstract | Display for All Results]

Molecule of the Month: Myoglobin, PDB Pioneers

Search Hit: Classification: **MYOGLOBIN** (CARBONMONOXY)

RCSB Protein Data Bank

www.rcsb.org/pdb/explore/explore.do?structureId=1MBN

Summary | PDB-101 | 3D View | Sequence | Annotations | Seq. Similarity | 3D Similarity | Literature | Biol. & Chem. | Methods | Geometry | Links

The stereochemistry of the protein myoglobin **1MBN**

Display Files
Download Files
Share this Page

DOI:10.2210/pdb1mbn/pdb

Primary Citation

The Stereochemistry of the Protein Myoglobin

Watson, H.C.

Journal: (1969) Prog.Stereochem. **4:** 299

PubMed ID is not available

Molecular Description

Classification: Oxygen Storage
Structure Weight: 17868.60

Molecule: MYOGLOBIN
Polymer: 1 Type: protein Length: 153
Chains: A
Organism Physeter catodon
Gene Name MB
UniProtKB: Protein Feature View | Search PDB | P02185

P02185
Molec. Processing Myoglobin
UP Sites
SCOP domains Globin-like
PDB Sites
Secstruc
1MBN.A

Biological Assembly

3D View More Images...

No symmetry
Stoichiometry: **Monomer**
Biological assembly 1 assigned by authors

Downloadable viewers:
Simple Viewer Protein Workshop
Kiosk Viewer

MyPDB Personal Annotations

To save personal annotations, please

PDB-101
Structural View of Biology
Understanding PDB Data
Molecule of the Month
Educational Resources
Author Profiles

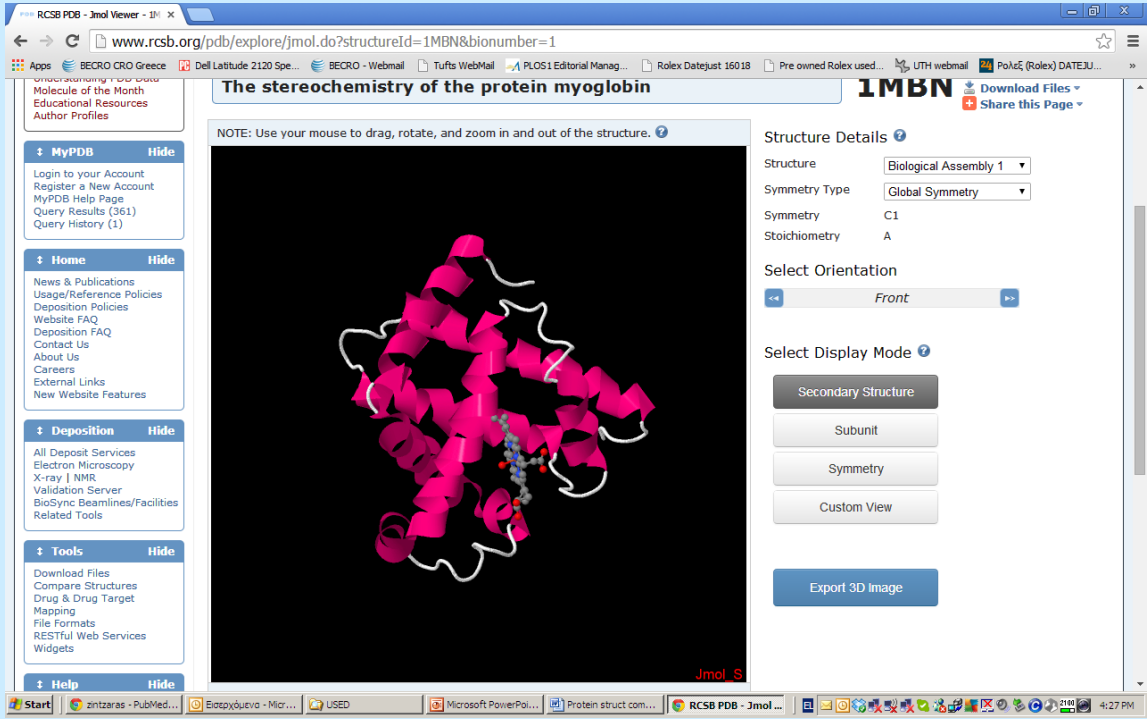
MyPDB
Login to your Account
Register a New Account
MyPDB Help Page
Query Results (361)
Query History (1)

Home
News & Publications
Usage/Reference Policies
Deposition Policies
Website FAQ
Deposition FAQ
Contact Us
About Us
Careers
External Links
New Website Features

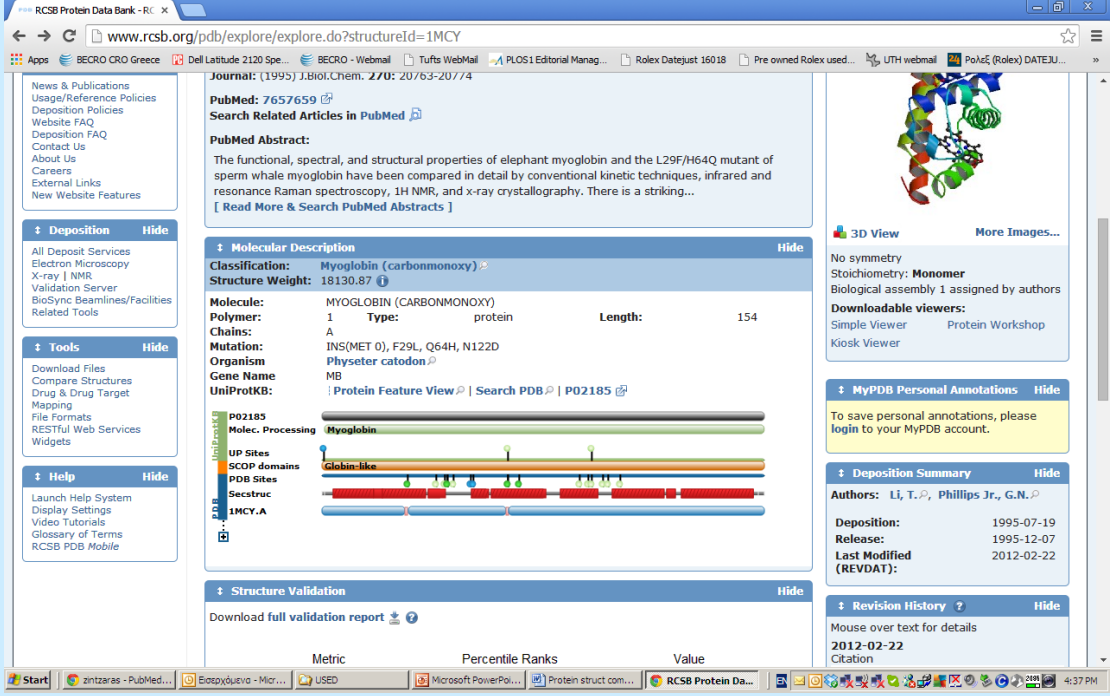
Deposition
All Deposit Services
Electron Microscopy
X-ray | NMR
Validation Server
BioSync Beamlines/Facilities
Related Tools

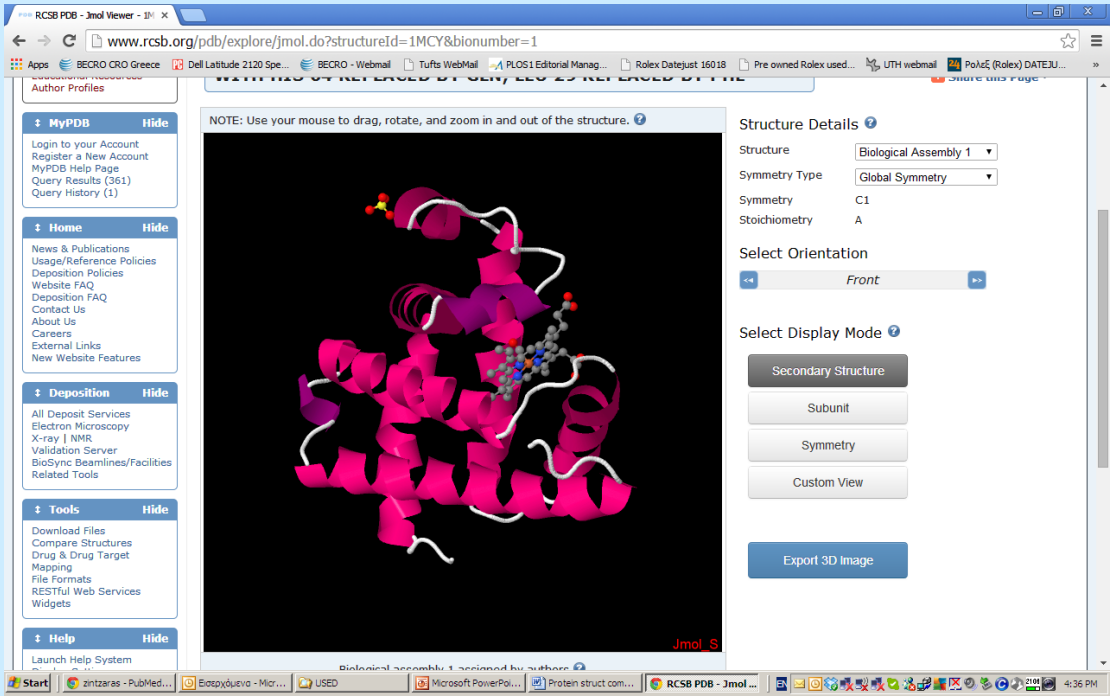
Tools
Download Files
Compare Structures
Drug & Drug Target
Mapping

Πατώντας εν συνεχεία στο link ‘3D View’, που βρίσκεται στα αριστερά της σελίδας μπορούμε να δούμε μία τρισδιάστατη εικόνα της πρωτεΐνης, όπου διακρίνονται και οι δευτεροταγείς δομές της.



Πατώντας στον κωδικό πρωτεΐνης ‘1MCY’, και μετά στο ‘3D View’ μπορούμε να δούμε πληροφορίες για μία άλλη myoglobin. Αν συγκρίνουμε τις εικόνες για τις δύο αυτές myoglobins (την 1MBN και την 1MCY) παρατηρούμε ότι παρουσιάζουν πολλές ομοιότητες ως προς τη δομή τους.

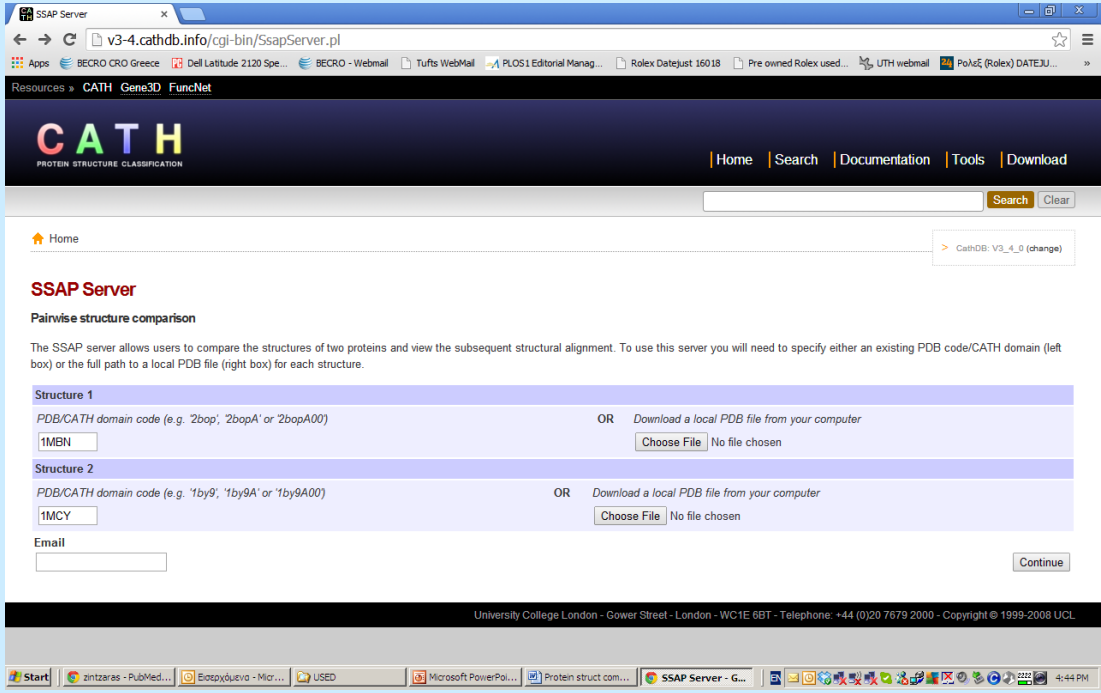




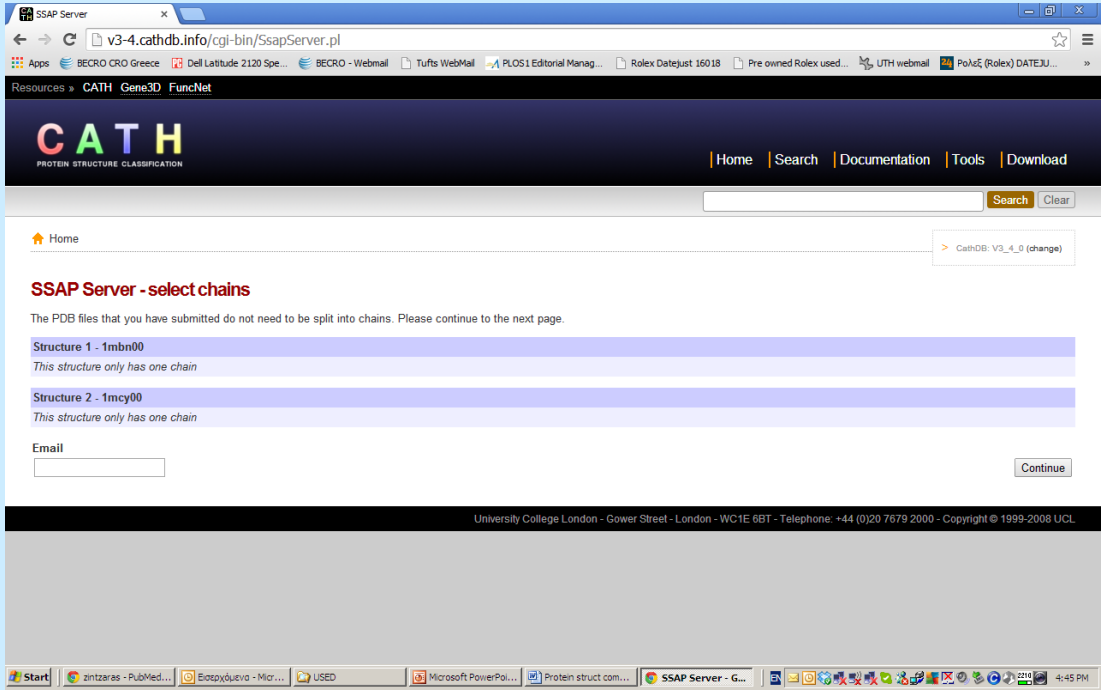
Για να συγκρίνουμε τώρα τις δομές των δύο πρωτεϊνών 1MBN και την 1MCY χρησιμοποιώντας τη μέθοδο SSAP, πληκτρολογήστε στον Internet Explorer τη διεύθυνση: www.cathdb.info/cgi-bin/SsapServer.pl

.

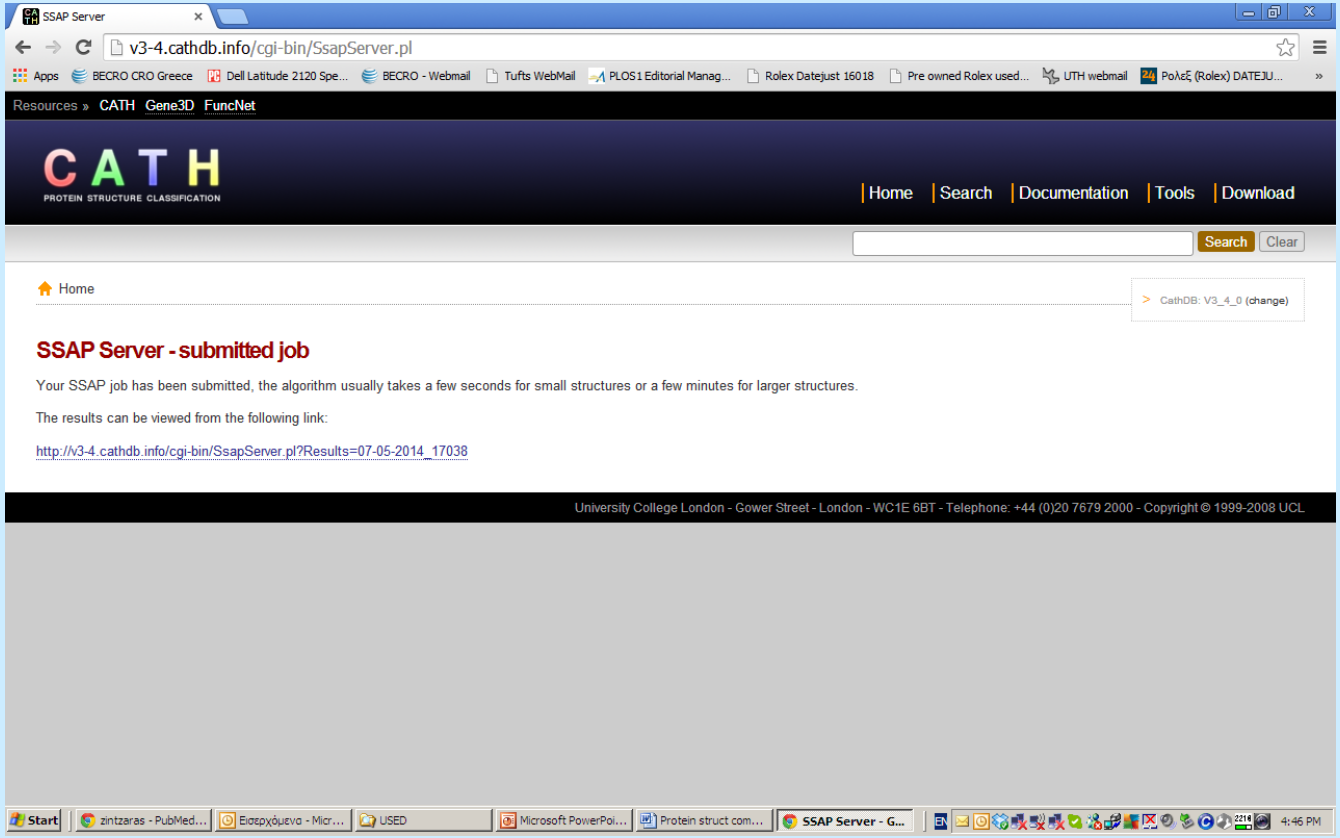
Στο κενό πεδίο κάτω από τον τίτλο ‘Structure 1’ πληκτρολογήστε 1MBN και στο πεδίο κάτω από το ‘Structure 2’ πληκτρολογήστε 1MCY. Έπειτα πατήστε ‘Continue’..



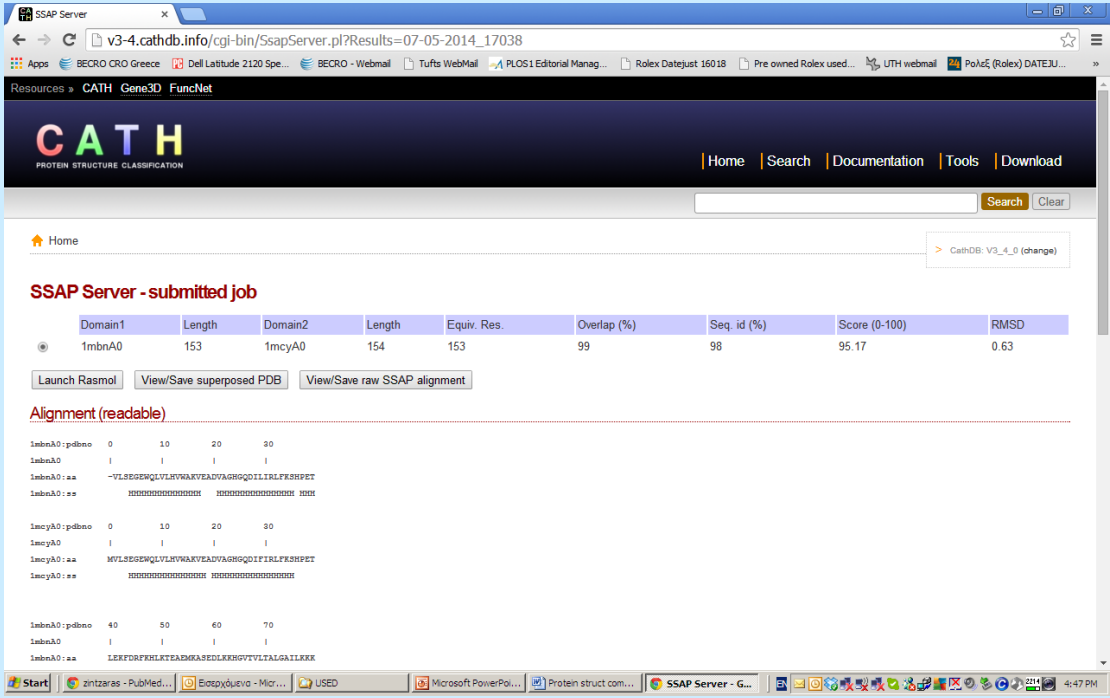
Παρουσιάζεται η δυνατότητα επιλογής για τις πρωτεϊνικές αλυσίδες προς σύγκριση, ωστόσο στη συγκεκριμένη περίπτωση του παραδείγματός μας, οι myoglobins έχουν μόνο μία αλυσίδα.

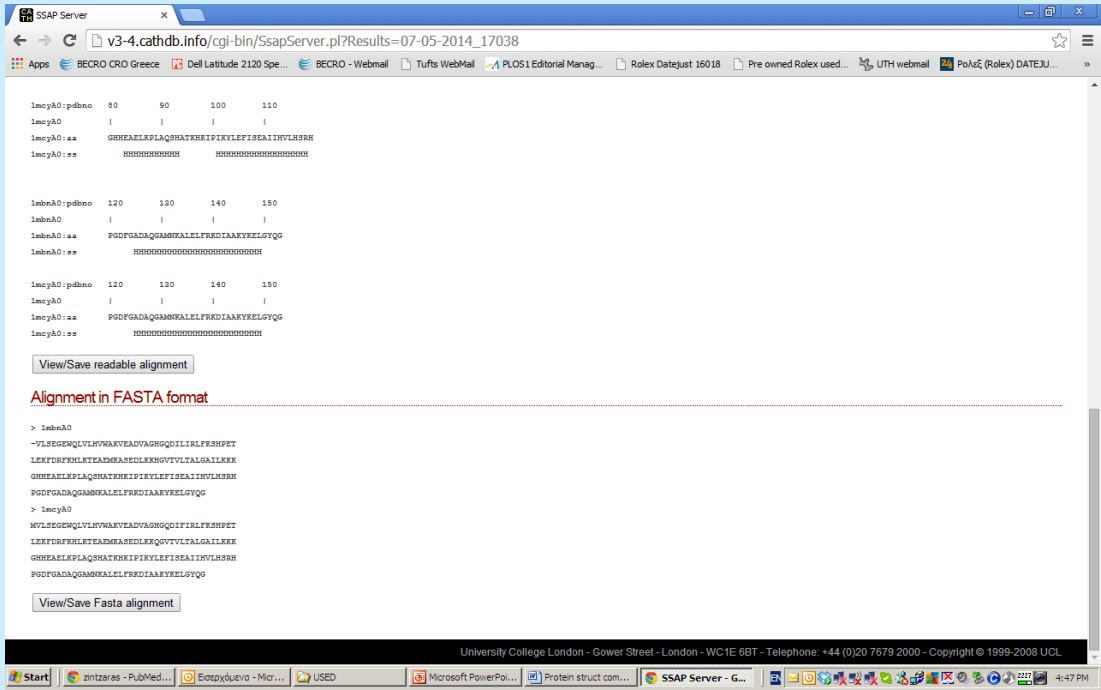


Για να δείτε τα αποτελέσματα πατήστε τη διεύθυνση που εμφανίζεται στο επόμενο παράθυρο.



Εκεί παρατίθεται με το RMSD η αντιστοίχηση των δομών για τις δύο πρωτεΐνες, η βαθμολογία ταύτισης και η ομοιότητα των αλληλουχιών. Επιπλέον, παρουσιάζεται ενδεικτικά και η δευτεροταγής δομή (secondary structure – ss).



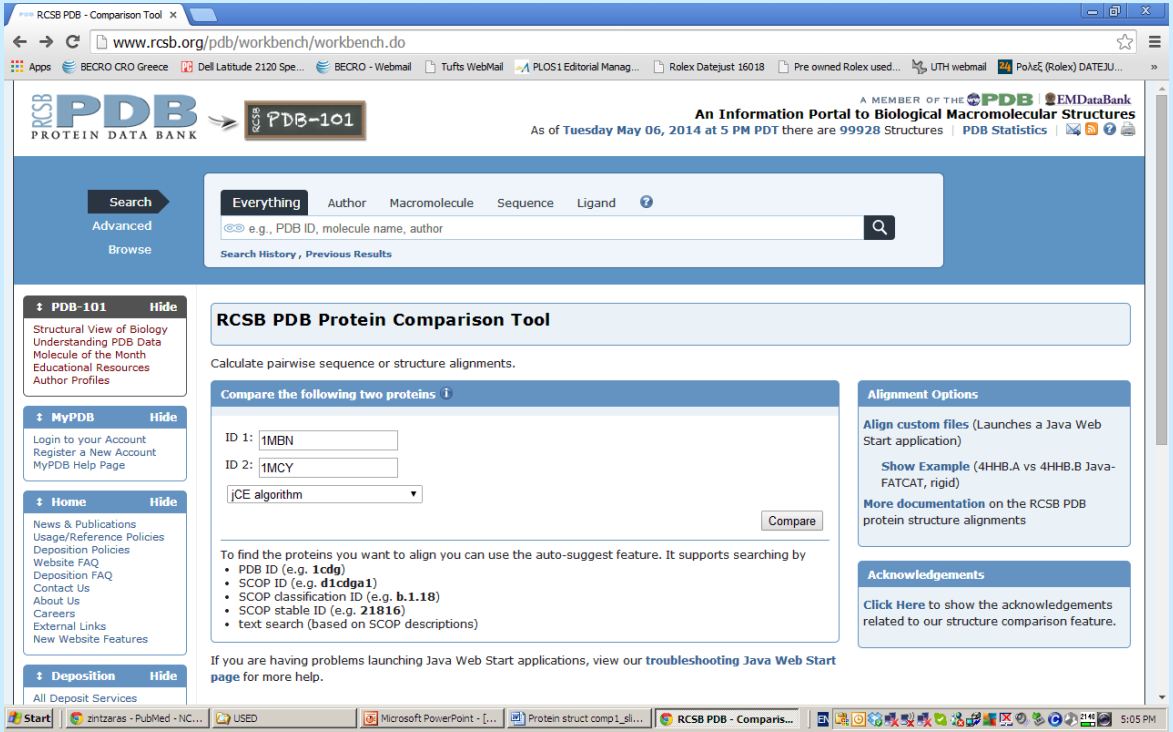
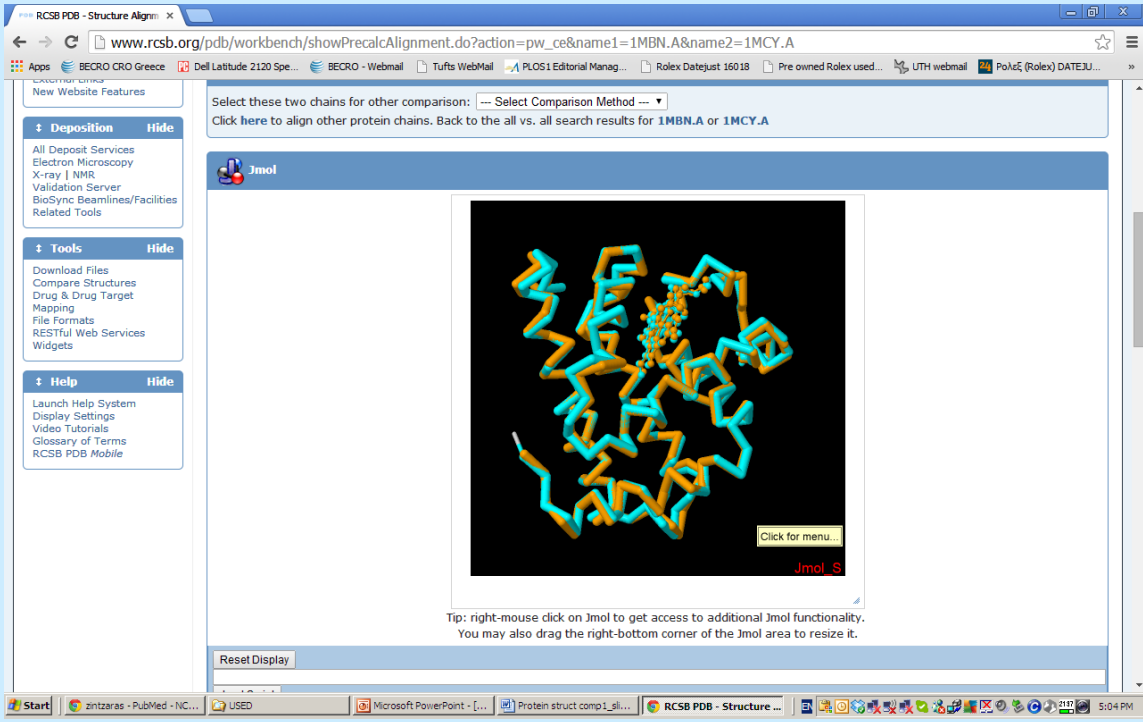
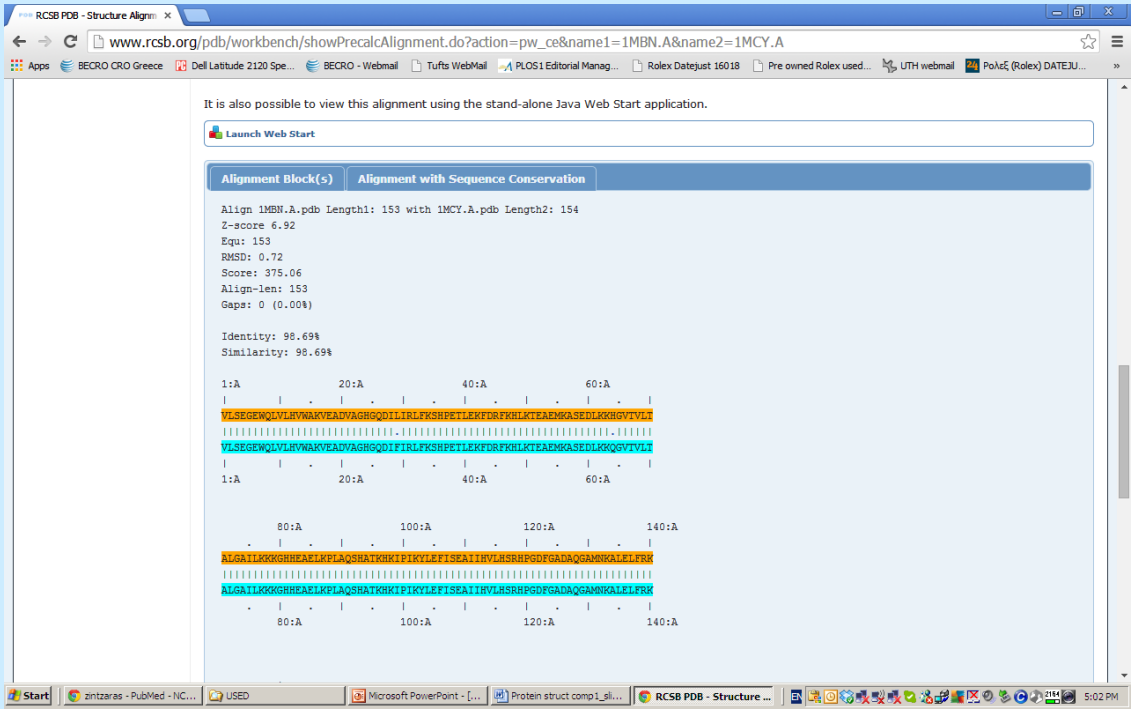


Συνδυασμός ενδομοριακών και διαμοριακών μεθόδων σύγκρισης

Μία άλλη μέθοδος για τη σύγκριση δομών πρωτεϊνών είναι η μέθοδος CE (Combinatorial Extension – Συνδυαστική Επέκταση), η οποία συνδυάζει τις ενδομοριακές και τις διαμοριακές μεθόδους (για λεπτομέρειες σχετικά με την μέθοδο CE συμβουλευτείτε την βοήθεια που ακολουθεί, ή το άρθρο του Tsigelny, Prot. Sci. 2000).

Για να επισκεφθείτε την ιστοσελίδα με πρόσβαση στη βάση δεδομένων για τη CE, ανοίξτε τον Internet Explorer και πληκτρολογήστε τη διεύθυνση: http://www.rcsb.org/pdb/workbench/workbench.do.

Για τη σύγκριση των πρωτεϊνών 1MBN και 1MCY, πληκτρολογήστε στο πεδίο της ID1: 1MBN και στο αντίστοιχο πεδίο για την ID2: 1MCY. Επίλεξτε “jCE algorithm” και έπειτα πατήστε το “Compare’ και θα προκύψουν τα αποτελέσματα της αντιστοίχησης, συμπεριλαμβανομένου του RMSD και της βαθμολογίας για την ομοιότητα των αλληλουχιών.



The alignment results are provided, including RMSD and sequence similarity score.