

Bazy Danych i Usługi Sieciowe

Bioinformatyczne bazy danych i usługi sieciowe

Paweł Daniluk

Wydział Fizyki

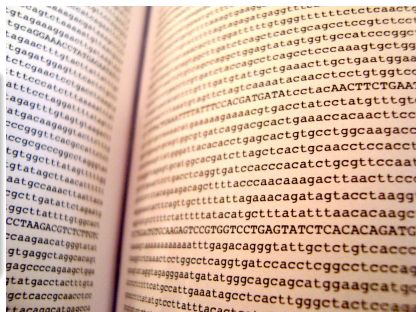
Jesień 2011



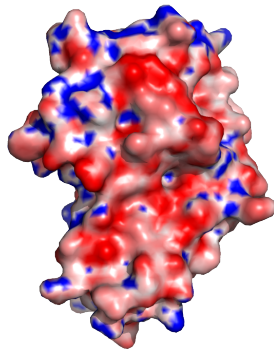
Zagadnienia

- katalogowanie informacji biologicznych
- analiza sekwencji DNA
- analiza sekwencji genomów, porównywanie genomów
- ustalanie ewolucyjnych relacji pomiędzy zbiorami sekwencji / organizmów (drzewa filogenetyczne)
- genotypowane
- analiza ekspresji genów
- analiza sekwencji białek – proteomika
- katalogowanie funkcji genów/białek, analiza dróg metabolicznych oraz szlaków sygnałowych
- modelowanie układów biologicznych
- wirtualne dokowanie (ang. virtual docking)
- komputery DNA
- morfometria / analiza obrazu

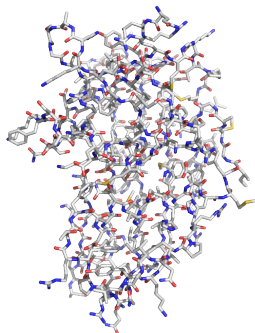
- sekwencje (nukleotydów, reszt aminokwasowych)



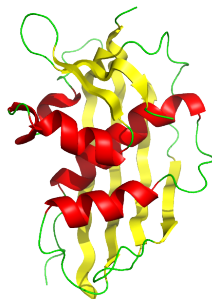
- sekwencje (nukleotydów, reszt aminokwasowych)
- struktury przestrzenne molekuł



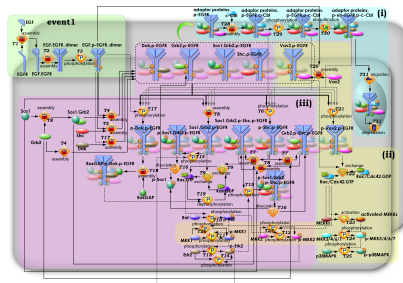
- sekwencje (nukleotydów, reszt aminokwasowych)
- struktury przestrzenne molekuł



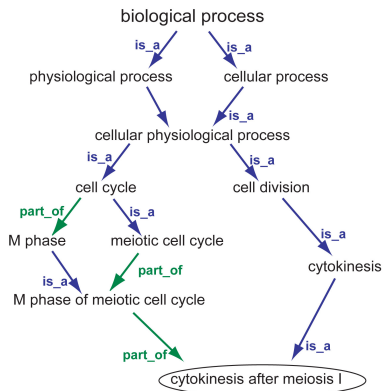
- sekwencje (nukleotydów, reszt aminokwasowych)
- struktury przestrzenne molekuł



- sekwencje (nukleotydów, reszt aminokwasowych)
- struktury przestrzenne molekuł
- sieci



- sekwencje (nukleotydów, reszt aminokwasowych)
- struktury przestrzenne molekuł
- sieci
- ontologie



Reprezentacja danych

- pliki tekstowe
- dane semi-strukturalne (XML)
- tabele

Reprezentacja danych

- pliki tekstowe
- dane semi-strukturalne (XML)
- tabele

FASTA - sekwencja reszt aminokwasowych

```
>gi|5524211|gb|AAD44166.1| cytochrome b [Elephas maximus maximus]
LCLYTHIGRNIYYGSYLSETWNTGIMLLITMATAFMGYVLPWGQMSFWGATVITNLFSAIPYIGTNLV
EWIWGGFSVDKATLNRFFAFHFILPFTMVALAGVHLTFLHETGSNNPLGLTSDSDKIPFHPYYTIKDFLG
LLILILLLLLLLALLSPDMLGDPDNHMPADPLNTPHLIKPEWYFLFAYAILRSPVNKLGGLALFLSIVIL
GLMPFLHTSKHRSMMLRPLSQALFWTLTMDLLTLTWIGSQPVEYPYTIIGQMASILYFSIILAFPLIAGX
IENY
```

Reprezentacja danych c.d.

PDB - struktura przestrzenna białka

```
HEADER      EXTRACELLULAR MATRIX                22-JAN-98   1A3I
TITLE       X-RAY CRYSTALLOGRAPHIC DETERMINATION OF A COLLAGEN-LIKE
TITLE       2 PEPTIDE WITH THE REPEATING SEQUENCE (PRO-PRO-GLY)
...
EXPDTA      X-RAY DIFFRACTION
AUTHOR      R.Z.KRAMER,L.VITAGLIANO,J.BELLA,R.BERISIO,L.MAZZARELLA,
AUTHOR      2 B.BRODSKY,A.ZAGARI,H.M.BERMAN
...
REMARK 350 BIOMOLECULE: 1
REMARK 350 APPLY THE FOLLOWING TO CHAINS: A, B, C
REMARK 350  BIOMT1   1  1.000000  0.000000  0.000000          0.00000
REMARK 350  BIOMT2   1  0.000000  1.000000  0.000000          0.00000
...
SEQRES     1 A      9  PRO PRO GLY PRO PRO GLY PRO PRO GLY
SEQRES     1 B      6  PRO PRO GLY PRO PRO GLY
SEQRES     1 C      6  PRO PRO GLY PRO PRO GLY
...
ATOM       1  N      PRO A   1          8.316  21.206  21.530  1.00 17.44      N
ATOM       2  CA     PRO A   1          7.608  20.729  20.336  1.00 17.44      C
ATOM       3  C      PRO A   1          8.487  20.707  19.092  1.00 17.44      C
ATOM       4  O      PRO A   1          9.466  21.457  19.005  1.00 17.44      O
ATOM       5  CB     PRO A   1          6.460  21.723  20.211  1.00 22.26      C
...
HETATM    130  C      ACY     401         3.682  22.541  11.236  1.00 21.19      C
HETATM    131  O      ACY     401         2.807  23.097  10.553  1.00 21.19      O
HETATM    132  OXT   ACY     401         4.306  23.101  12.291  1.00 21.19      O
...
```

PDBML - struktura przestrzenna białka

(a) Example of a fully marked-up PDBML atom record

```
<PDBx:atom_siteCategory>
  <PDBx:atom_site id="1">
    <PDBx:group_PDB>ATOM</PDBx:group_PDB>
    <PDBx:type_symbol>N</PDBx:type_symbol>
    <PDBx:label_atom_id>N</PDBx:label_atom_id>
    <PDBx:label_comp_id>ASP</PDBx:label_comp_id>
    <PDBx:label_asym_id>A</PDBx:label_asym_id>
    <PDBx:label_entity_id>1</PDBx:label_entity_id>
    <PDBx:label_seq_id>1</PDBx:label_seq_id>
    <PDBx:Cartn_x>23.482</PDBx:Cartn_x>
    <PDBx:Cartn_y>-0.621</PDBx:Cartn_y>
    <PDBx:Cartn_z>-1.419</PDBx:Cartn_z>
    <PDBx:occupancy>1.00</PDBx:occupancy>
    <PDBx:B_iso_or_equiv>35.27</PDBx:B_iso_or_equiv>
    <PDBx:auth_seq_id>1</PDBx:auth_seq_id>
    <PDBx:auth_comp_id>ASP</PDBx:auth_comp_id>
    <PDBx:auth_asym_id>A</PDBx:auth_asym_id>
    <PDBx:auth_atom_id>N</PDBx:auth_atom_id>
    <PDBx:pdbx_PDB_model_num>1</PDBx:pdbx_PDB_model_num>
  </PDBx:atom_site>
</PDBx:atom_siteCategory>
```

(b) Example of the alternative simplified XML PDB atom record

```
<category_atom_record>
  <atom_record id="1">
    ATOM 1 A A 1 1 ? . ASP ASP N N N 23.482 -0.621 -1.419 1.00 35.27
  </atom_record>
</category_atom_record>
```

Usługi bioinformatyczne

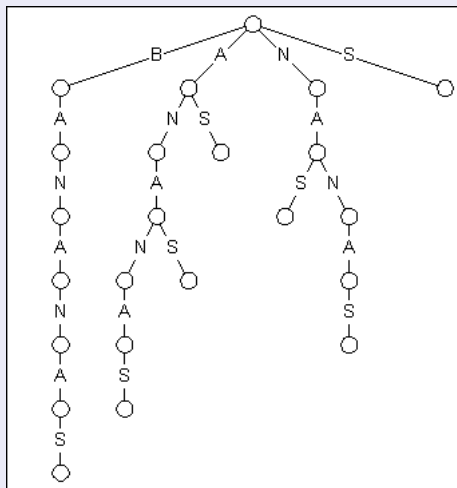
- porównywanie (sekwencji, struktur)
- wyszukiwanie w zbiorze danych
- analiza sekwencji
- mnóstwo innych

Bioinformatyczne bazy danych

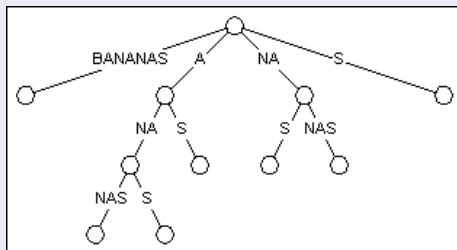
- duże ilości danych
- brak spójnego schematu
- skomplikowane kryteria wyszukiwania (często wykonuje się wstępne obliczenia)

Przeszukiwanie

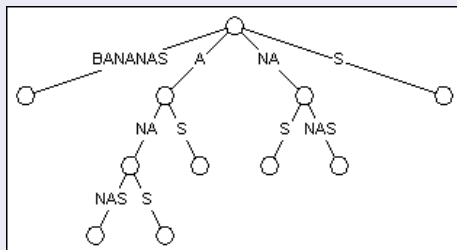
Drzewa sufiksowe



Drzewa sufiksowe



Drzewa sufiksowe



- Konieczne jest wstępne obliczenie.
- Możliwe jest szybkie stwierdzenie, czy istnieje sufiks zaczynający się od dowolnego słowa.
- Drzewo sufiksowe można zbudować dla wielu słów na raz.

RCSB PDB

http://www.rcsb.org/pdb/home/home.do

A MEMBER OF THE PDB PROTEIN DATA BANK

An Information Portal to Biological Macromolecular Structures
As of Tuesday Dec 27, 2011 at 4 PM PST there are 78237 Structures | PDB Statistics

All Categories Author Macromolecule Sequence Ligand

Search | All Categories: e.g. PDB ID, molecule name, author

Customize This Page

HyPDB Home Deposition Tools PDB-101 Help

Biological Macromolecular Resource

Full Description

Featured Molecules

Structural View of Biology

Molecule of the Month

Complex 1
Complex 1, also known as NADH:quinone oxidoreductase, performs the first step in respiratory electron transport, the process that creates much of the energy that powers our cells.

Full Article

Protein Structure Initiative Featured System

Making Lipopolysaccharide
Many bacteria surround themselves with a protective coat, to resist attack from antibiotics, predators, and the immune system. In gram-negative bacteria like Escherichia coli, this protective coat is built primarily of lipopolysaccharides (shown here from PDB entry 1JG1), hybrid molecules with a complex carbohydrate anchored to the cell membrane with an array of lipids.

Full Article | Archive | PSI Structural Biology Knowledgebase

Explore Archive

Organization Taxonomy Organism

File Method 3-Ray Resolutions

Release Date Polymer Type

Keywords Classification SCOP Classification

Show All

Latest Structures

3uht - HBI (L36V) deoxy
Ren, Z., Sreer, V., Knapp, J.E., Royer, J., W.E.

Cooperative macromolecular device revealed by meta-analysis of static and time-resolved structures.
(2011) Proc Natl Acad Sci USA

Read Full Abstract

View in 3D (Jmol)

Check out all of the latest structures released.

New user? Try the browser Compatibility Check and information on Getting Started, or the Narrated Tutorial.

New Structures Latest Release New Structure Papers Unreleased Entries

New Features Webservice for Pre-released Sequences Latest Features Released Withds Release Archive

RCSB PDB News Weekly | Quarterly | Yearly

Happy Holidays

Best wishes for the new year, more

Notes for RIFMAC Users Looking for Ligands Paper Published Latest Website Release

wwPDB News X-ray Validation Task Force Report Published

DM Data Bank to join the PDB archive

POB40 Symposium Update Full wwPDB News Statement on Restriction of PDB Entries

FTP Archives Current PDB FTP Archive: ftp.wwpdb.org Yearly FTP Snapshot: ftp://anapshot.wwpdb.org

1 RCSB PDB - 2WY4 Structure Summary

http://www.rcsb.org/pdb/explore/structure/2WY4

Summary Sequence Annotations Seq. Entry 3D Entry Literature Ref. & Chem. Methods Geometry Links

STRUCTURE OF BACTERIAL GLOBIN FROM CAMPYLOBACTER JEJUNI AT 1.35 Å RESOLUTION

2WY4

DOI:10.2210/pdb/2wy4/jdb

Primary Citation

The single-domain globin from the pathogenic bacterium *Campylobacter jejuni* novel D-helix conformation, proximal hydrogen bonding that influences ligand binding, and peroxidase-like redox properties.

Shepherd, M., Barynin, V.V., Lu, C., Bernhardt, P.V., Wu, G., Yeh, S.R., Ogawa, T., Sedelnikova, S.E., Rice, D.W., Wilson, J.L., Paoletti, R.K.,

Journal: (2010) *J.Biol.Chem.* **285**: 12747-12754

PubMed: 20164176 [?] [View Article](#)

PubMedCentral: PMC2857070 [?] [View Article](#)

DOI: 10.1074/jbc.M109.084509 [?] [View Article](#)

[Search Related Articles in PubMed](#) [?]

PubMed Abstract:

The food-borne pathogen *Campylobacter jejuni* possesses a single-domain globin (CjgB) whose role in detoxifying nitric oxide has been unequivocally demonstrated through genetic and molecular approaches. The x-ray structure of cyanide-bound CjgB has been solved to a resolution of 1.35 Å...

[[Read More & Search PubMed Abstracts](#)]

Monomeric Description

Classification: Cytochrome Transport

Structure Weight: 16745.52

Molecule: SINGLE DOMAIN HAEMGLOBIN

Polymer: 1 Type: protein Length: 140

Chain: A

Details: CN

Organism: *Campylobacter jejuni* [?]

UniProtKB: Q9P842 [?]

Source

Polymer: 1

Scientific Name: *Campylobacter* Taxonomy: *Escherichia coli* [?]

Name: jejuni [?]

Expression System:

Ligand Chemical Component

Identifier	Formula	Name	Interactions
CYN	C N	CYANIDE ION	Ligand Explorer
HEM	C ₃₄ H ₃₂ Fe N ₄ O ₄	PROTOPORPH	Ligand Explorer

External Domain Annotations

- CATH Classification: v3.4.0: 1 Domain - data from CATH [?]
- PFAM Classification: 1 Domain - data from PFAM [?]
- GO Terms: 5 Terms - data from GO [?]

Structural Biology Knowledgebase Data

Information from the Structural Biology Knowledgebase [?]

- Models from the Protein Model Portal: 1 model [?]
- Protein Targets from TargetDB: 0 targets [?]
- Nature Biological Annotations: 0 annotations [?]
- Related Clones from PSI-BioSIS Materials Repository: 0 clones [?]

Biological Assembly

More Images...

[View in Small](#) [Simple Viewer](#)
[Other Viewers](#) [Protein Workshop](#)

Biological assembly 1 generated by PISA (software)

Deposition Summary

Author: Barynin, V.V., Sedelnikova, S.E., Shepherd, M., Wu, G., Paoletti, R.K., Rice, D.W.,

Deposition: 2009-11-11

Release: 2010-02-16

Last Modified (REVDAT): 2010-04-28

Experimental Details

Method: X-RAY DIFFRACTION

Exp. Date:

Structure Factors

ESF [?]

Resolution [Å]: 1.35

R-Value: 0.180 (obs.)

R-Free: 0.209

Space Group: P 2₁

Unit Cell:

Length [Å]: Angles [°]

a = 39.55 a = 90.00

b = 37.98 b = 102.68

c = 43.76 c = 90.00

NCBI Blast Protein Sequence (78 letters)

http://blast.ncbi.nlm.nih.gov/Blast.cgi

BLAST® Basic Local Alignment Search Tool

Home Recent Results Saved Strategies Help

NCBI BLAST! [Help](#) [Formatting Results - G4EYD0H13](#)

[Edit and Resubmit](#) [Save Search Strategies](#) [Formatting options](#) [Download](#)

PSI Blast Iteration 1

Protein Sequence (78 letters)

Query ID: K156580
 Description: None
 Molecule type: amino acid
 Query Length: 78

Database Name: nr
 Description: All non-redundant GenBank CDS translations + PDB + SwissProt + TrEMBL excluding environmental samples from WGS projects
 Program: BLASTP 2.2.30+ -i330000

Other reports: [Search Summary](#) [Taxonomy reports](#) [Distance tree of results](#) [Related Structures](#) [Multiple alignment](#)

Graphic Summary

Show Conserved Domains

Putative conserved domains have been detected, click on the image below for detailed results.

Query seq. Superfamily Multi-domain

Distribution of 112 Blast Hits on the Query Sequence

Color key for alignment scores

Score Range	Color
>40	Black
40-60	Blue
60-80	Green
80-100	Yellow
>=100	Red

Legend for links to other resources: UniGene, Gene, Structure, Map Viewer, PubChem BioAssay

Legend for alignment status:

- - alignment score below the threshold on this previous iteration
- - alignment was checked on the previous iteration

Run PSI-Blast iteration 2 with max

Sequences producing significant alignments with E-value BETTER than threshold

NCBI Blast Protein Sequence (78 letters)

Legend for links to other resources: [UniGene](#) [Gene](#) [Structure](#) [Map Viewer](#) [PubChem BioAssay](#)

Alignment score below threshold or on previous iteration
 Alignment was checked on the previous iteration
 Run PSI-Blast iteration 2 with max 500

Sequences producing significant alignments with E-value BETTER than threshold

Accession	Description	Max score	Total score	Query cover	E-value	Max ident	Links
M1301_6	Chain A, NK-Lysin From <i>Fig. Nere, 2D</i> Structures	153	159	100%	2e-50	100%	G
M029275_1	Neckline: Full-Antimicrobial peptide NK-lysin; Short-NKL1	153	145	100%	2e-44	91%	G
MF01202932.1	PREDICTED: antimicrobial peptide NK-lysin [Sua acrofla]	143	145	100%	6e-44	91%	G M
MF01022309.2	antimicrobial peptide NK-lysin precursor [Staurus cabaflus] >	103	109	100%	8e-30	69%	G M
MF0101222.2	antimicrobial peptide NK-lysin [Staurus cabaflus]	103	109	100%	1e-29	69%	G M
MF01028732.1	granulysin precursor [Dvia arles] >gi AM807873.1 granul	103	105	94%	3e-28	67%	G M
M0217339_2	granulysin [Stubalus bubalis]	98.6	94.4	98%	1e-24	56%	G
MF020032.1	granulysin/NK-lysin-like protein [Bos taurus]	98.2	94.7	94%	1e-24	59%	G M
MF0102861.1	granulysin precursor [Bos taurus] >gi BM424279.3 Granu	95.1	95.1	94%	3e-24	59%	G M
MF020032.1	granulysin/NK-lysin-like protein [Bos taurus]	86.3	86.3	96%	2e-31	55%	G
M02145177.1	Granulysin [Bos taurus]	82.6	82.4	94%	3e-19	52%	G
MF010338293.1	PREDICTED: antimicrobial peptide NK-lysin-like [Bos taurus]	81.6	81.6	96%	7e-19	51%	G
MF010291451.3	PREDICTED: antimicrobial peptide NK-lysin [Bos taurus] >	81.3	81.3	96%	1e-18	51%	G M
MF020032.1	hypothetical protein LOC018323 [Bos taurus]	78.6	78.6	96%	8e-18	49%	G
MF010104043.1	uncharacterized protein LOC018323 precursor [Bos taurus]	78.6	78.6	96%	9e-18	49%	G M
MF0103288937.1	PREDICTED: hypothetical protein LOC1004602334 [Nomase	67.8	60.8	94%	5e-10	37%	G M
MF0103811865.1	PREDICTED: granulysin-like [Pongia abelii]	52.8	57.8	98%	9e-10	37%	G
MF01033392.2	PREDICTED: hypothetical protein LOC049370 isoform 2 [P	52.1	55.1	97%	1e-08	34%	G M
MF010168372.1	PREDICTED: hypothetical protein LOC049370 isoform 1 [P	52.1	55.1	97%	1e-08	34%	G M
MF0103305128.1	PREDICTED: hypothetical protein LOC049370 [Pan troglody	54.7	54.7	97%	3e-08	34%	G
MF0102757576.1	PREDICTED: hypothetical protein LOC100400940 [Callithrix	56.2	56.2	98%	2e-08	35%	G
MF010342253.1	PREDICTED: antimicrobial peptide NK-lysin-like [Bos tauru	54.2	54.7	97%	3e-08	32%	G M
M0212996.1	NK-lysin type 2 [Jastura punctulata]	53.5	53.5	93%	4e-08	32%	G
MF010118747.1	NK-lysin type 2 precursor [Stalulus punctulata] >gi ABC7	53.1	53.5	93%	4e-08	32%	G M
MF010342241.4	PREDICTED: uncharacterized protein LOC10068082 [Canis	53.1	55.1	97%	5e-08	36%	G M
MF010118732.1	NK-lysin type 1 precursor [Stalulus punctulata] >gi AM176	51.2	51.2	94%	4e-07	31%	G M
MF010273518.1	GK13189 [Drosophila williamsi] >gi DQW4504.1 GK131	50.8	50.8	93%	2e-07	34%	G M
MF010281451.1	NK-lysin precursor [Stalulus punctulata] >gi AM60870.1 NK-ly	48.9	48.9	98%	3e-05	25%	G M
MF01035615.2	granulysin isoform S19 [Homo sapiens]	48.5	48.5	91%	3e-06	33%	G M
MF01036424.2	granulysin isoform NGS precursor [Homo sapiens] >gi P2	48.5	48.5	91%	4e-06	33%	G M
M02045493.1	unlabeled protein product [Homo sapiens]	48.5	48.5	91%	5e-05	33%	G
MF01026215.1	unlabeled protein product [Homo sapiens]	48.5	48.6	91%	2e-05	33%	G M
MF01018184.1	granulysin, isoform CRA_3 [Homo sapiens]	48.5	48.6	91%	1e-05	33%	G
MF0103214074.1	PREDICTED: ascocin-C-like [Melospiza gallopavo]	47.0	47.0	98%	2e-05	25%	G M
M02045493.1	NGS protein [Homo sapiens] >gi EAW94843.1 granulysin	48.5	48.6	91%	2e-05	33%	G
MF01018184.1	granulysin, isoform CRA_c [Homo sapiens]	48.5	48.6	91%	5e-05	33%	G
MF0101187232.1	NK-lysin type 2 precursor [Stalulus punctulata] >gi ABC7	53.1	45.1	96%	7e-05	27%	G M
MF010134552.1	antimicrobial peptide NK-lysin precursor [Stalulus punctulata]	53.1	45.1	94%	8e-05	30%	G M
M0212996.1	NK-lysin type 2 [Stalulus punctulata]	43.3	43.5	98%	3e-04	27%	G
M131_6	Chain A, Granulysin From Human Cytolytic T Lymphocytes	42.6	42.4	88%	3e-04	33%	G
MF010202412.1	GL13982 [Drosophila paramela] >gi EDW39229.1 GL139	46.7	44.7	93%	3e-04	31%	G M
MF010312588.1	CA11372 [Drosophila pseudoobscura pseudoobscura] >gi E	44.3	44.3	93%	4e-04	31%	G M
MF010208452.1	GL13922 [Drosophila yakuba] >gi P4059102.1 GL1392	43.3	43.9	93%	6e-04	29%	G M
MF01025846.1	PREDICTED: antimicrobial peptide NK-lysin-like [Canis lupus	43.1	43.1	94%	7e-04	28%	G M
MF01020517.1	PREDICTED: antimicrobial peptide NK-lysin-like [Canis lupus	42.7	42.7	96%	7e-04	28%	G M
MF010364064.1	GF23306 [Drosophila americana] >gi EDV30460.1 GF23	43.3	43.5	93%	7e-04	30%	G M
MF0102709.1	uncharacterized protein LOC331953 precursor [Danio rerio]	42.6	42.6	93%	0.001	24%	G M
MF010293125.1	PREDICTED: antimicrobial peptide NK-lysin-like [Alburnops	41.6	41.6	67%	0.002	34%	G M
MF01022281.1	hypothetical protein GSK_05326, partial [Descaia mullata]	40.8	40.8	70%	0.002	23%	G
MF010200363.1	PREDICTED: antimicrobial peptide NK-lysin-like [Danio rerio]	41.2	41.2	93%	0.002	23%	G M

SCOP: Root: scop

http://scop.mrc-lmb.cam.ac.uk/scop/data/scop.b.html

Google

Apple Yahoo! Google Maps YouTube Wikipedia News (513) Popular Strzelaj

Structural Classification of Proteins

Root: scop

Classes:

1. [All alpha proteins](#) [46456] (284)
2. [All beta proteins](#) [48724] (174)
3. [Alpha and beta proteins \(a/b\)](#) [51349] (147)
Mainly parallel beta sheets (beta-alpha-beta units)
4. [Alpha and beta proteins \(a+b\)](#) [53931] (376)
Mainly antiparallel beta sheets (segregated alpha and beta regions)
5. [Multi-domain proteins \(alpha and beta\)](#) [56572] (66)
Folds consisting of two or more domains belonging to different classes
6. [Membrane and cell surface proteins and peptides](#) [56835] (58)
Does not include proteins in the immune system
7. [Small proteins](#) [56992] (90)
Usually dominated by metal ligand, heme, and/or disulfide bridges
8. [Coiled coil proteins](#) [57942] (7)
Not a true class
9. [Low resolution protein structures](#) [58117] (26)
Not a true class
10. [Peptides](#) [58231] (121)
Peptides and fragments. Not a true class
11. [Designed proteins](#) [58788] (44)
Experimental structures of proteins with essentially non-natural sequences. Not a true class

Enter [search](#) key:

Generated from scop database 1.75 with scopm 1.101 on Wed Jun 3 10:42:06 2009
Copyright © 1994-2009 The scop authors / scop@mrc-lmb.cam.ac.uk

Go to "http://scop.mrc-lmb.cam.ac.uk/scop/data/scop.b.c.html"

Większość dobrych czasopism wymaga, aby algorytmy opisywane w publikowanych artykułach miały publicznie dostępne implementacje.

Typowy scenariusz

- 1 Użytkownik otwiera stronę z formularzem.
- 2 Użytkownik podaje parametry zapytania, adres e-mail i wysyła zapytanie.
- 3 Zapytanie jest rejestrowane w systemie i umieszczane w kolejce.
- 4 Wyświetlana jest informacja o przyjęciu zapytania, jego identyfikatorze i ew. przybliżonym czasie oczekiwania.
- 5 Zapytanie jest wykonywane.
- 6 Przygotowywana jest strona z wynikiem.
- 7 Użytkownik otrzymuje e-mail z linkiem do wyniku.

Usługi bioinformatyczne – przykład

DEDAL on Essentia Proteomica server. - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://bioexploratorium.pl/EP/DEDAL/

DEDAL on Essentia Proteomica ...

Descriptor Structural Alignment

Mode: TS, descriptors with ≥ 3 segments + CTS refinement on all descriptors

Max. sequence offset:

Max. sequence swaps:

First structure

Insert structure code (PDB or ASTRAL):

or upload structure file:

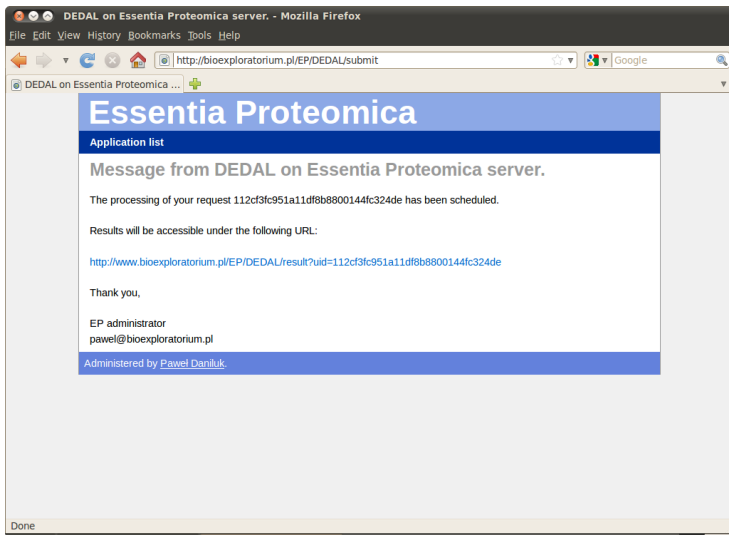
Second structure

Insert structure code (PDB or ASTRAL):

or upload structure file:

Administered by Paweł Daniluk.

Usługi bioinformatyczne – przykład



The screenshot shows a Mozilla Firefox browser window with the address bar containing `http://bioexploratorium.pl/EP/DEDAL/submit`. The page title is "DEDAL on Essentia Proteomica server." The main content area features a blue header with the text "Essentia Proteomica" and a sub-header "Application list". Below this, a white box contains the following text:

Message from DEDAL on Essentia Proteomica server.

The processing of your request 112cf3fc951a11df8b8800144fc324de has been scheduled.

Results will be accessible under the following URL:

<http://www.bioexploratorium.pl/EP/DEDAL/result?uid=112cf3fc951a11df8b8800144fc324de>

Thank you,

EP administrator
pawel@bioexploratorium.pl

Administered by [Pawel Daniluk](#).

The status bar at the bottom of the browser window displays the word "Done".

Usługi bioinformatyczne – przykład

DEDAL on Essentia Proteomica server. - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://www.bioexploratorium.pl/EP/DEDAL/result?uid=112cf3fc951a11df8b8800144fc324de

Google

DEDAL on Essentia Proteomica ...

Essentia Proteomica

Application list

Results for BIAS request 112cf3fc951a11df8b8800144fc324de

There are no results yet, please come back later or wait for this page to reload.

Administered by [Paweł Daniluk](#).

Done

Usługi bioinformatyczne – przykład

DEDAL on Essentia Proteomics server. - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://www.bioexploratorium.pl/EP/DEDAL/result?uid=112cf3fc951a11df8b8800144fc324de

Essentia Proteomics

Application list

Results for BIAS request 112cf3fc951a11df8b8800144fc324de

Parameters used	
max. sequence swaps	-1
mode	6
max. sequence offset	10000
second structure code	d1qdma1
first structure code	d1nkla_

Descriptor Structural Alignment of 1nkla_ and 1qdma1

#	N. AA	scored AA	RMSD	Score	Tension	Alignment	Regions
Superposition							
Alignment 1	74	74	2.75	69.58	2.1	Ranges FASTA	Mol 1 Mol 2
Regions R1	74	74	2.75			PDB RasMOL script	
Alignment 2	71	71	3.3	65.62	2.32	Ranges FASTA	Mol 1 Mol 2
Regions R1	71	71	3.3			PDB RasMOL script	
Alignment 3	68	68	2.46	62.6	2.32	Ranges FASTA	Mol 1 Mol 2

Taverna Workbench

Podejście z formularzem jest przyjazne dla użytkownika, ale niezbyt użyteczne do automatyzacji.

Protokoły wywoływania usług i ich opisu

- SOAP (pierwotnie Simple Object Access Protocol)
- Web Services Description Language (WSDL, pronounced 'wiz-dul')

Możliwość łączenia wielu usług (dostępnych na różnych serwerach) w większe schematy.

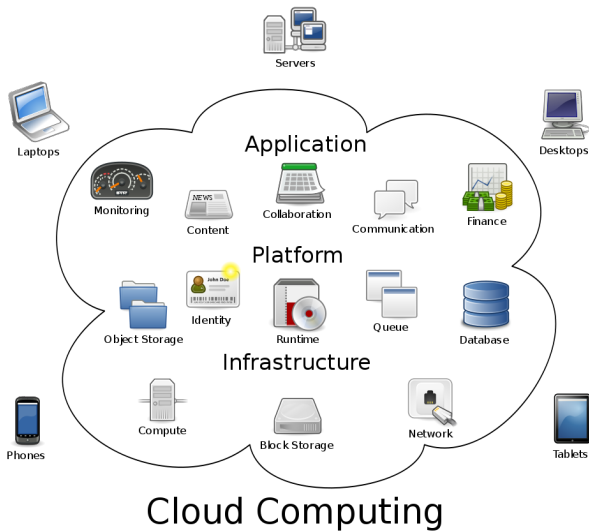
Taverna Workbench – c.d.

The screenshot displays the Taverna Workbench v1.7.1.0 interface. On the left, the 'Advanced model explorer' shows a workflow named 'Fetch Dragon images from BioMoby'. The workflow is composed of several processors: 'id', 'namespace', 'Object', 'getDragonSimpleAnnotatedImages', 'getUpegFromAnnotatedImage', 'Parse_Moby_Data_JPEGImage', 'Decode_base64_to_byte', and 'Parse_Moby_Data_SimpleAnnotatedJPEGImage'. A table below the processor list shows various properties for each processor.

Processor	0	0	1	1
id	0	0	1	1
namespace	0	0	1	1
Decode_base64_to_byte	0	0	1	1
getUpegFromAnnotatedImage	0	0	1	1
getDragonSimpleAnnotatedImages	0	0	1	1
Object	0	0	1	1
Parse_Moby_Data_JPEGImage	0	0	1	1
Parse_Moby_Data_SimpleAnnotatedJPEGImage	0	0	1	1

The main workspace shows a graphical workflow diagram. The workflow starts with 'id' and 'namespace' inputs leading to an 'Object' processor. This is followed by 'getDragonSimpleAnnotatedImages', then 'getUpegFromAnnotatedImage', and 'Parse_Moby_Data_JPEGImage'. The 'Parse_Moby_Data_JPEGImage' processor has two outputs: 'Decode_base64_to_byte' and 'Parse_Moby_Data_SimpleAnnotatedJPEGImage'. The 'Decode_base64_to_byte' processor outputs to 'images', and the 'Parse_Moby_Data_SimpleAnnotatedJPEGImage' processor outputs to 'annotations'. The 'images' and 'annotations' outputs are grouped under 'Workflow Outputs'.

Cloud computing



Warstwy

- Klient
- Usługi (SaaS – Software as a Service) – np. Google Apps, iCloud
- Platformy (PaaS – Platform as a Service) – np. Google App Engine, serwisy hostingowe
- Infrastruktura (IaaS – Infrastructure as a Service) – np. Amazon EC2
- Serwery

Cloud computing – wady i zalety

Zalety

- skalowalność
- niezależność od miejsca
- niższy koszt (efekt skali, inna kwalifikacja kosztu)
- łatwość konserwacji
- bezpieczeństwo
- daje większe możliwości użytkownikom końcowym

Cloud computing – wady i zalety

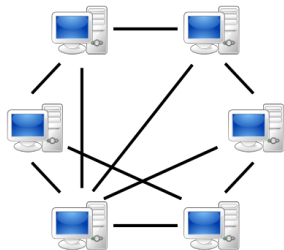
Zalety

- skalowalność
- niezależność od miejsca
- niższy koszt (efekt skali, inna kwalifikacja kosztu)
- łatwość konserwacji
- bezpieczeństwo
- daje większe możliwości użytkownikom końcowym

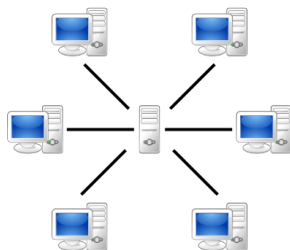
Wady

- brak kontroli nad danymi (data escrow)
- mniejsza możliwość dostosowania do szczególnych wymagań
- wielość standardów
- problemy prawne
- bezpieczeństwo

Peer to Peer (P2P)



vs.



Peer to Peer (P2P) – c.d.

P2P

- podejście czyste – wszystkie węzły są jednakowe
- podejście hybrydowe – superwęzły
- podejście scentralizowane

Peer to Peer (P2P) – c.d.

P2P

- podejście czyste – wszystkie węzły są jednakowe
- podejście hybrydowe – superwęzły
- podejście scentralizowane

Zalety

- lepsze wykorzystanie zasobów (przepustowości)
- łatwość konfiguracji
- niezawodność

Peer to Peer (P2P) – c.d.

P2P

- podejście czyste – wszystkie węzły są jednakowe
- podejście hybrydowe – superwęzły
- podejście scentralizowane

Zalety

- lepsze wykorzystanie zasobów (przepustowości)
- łatwość konfiguracji
- niezawodność

Wady

- niebezpieczeństwo “wstrzyknięcia” błędnych treści
- brak kontroli nad zawartością
- problemy prawne (cenzura internetu)
- blokowanie przez ISP (net neutrality?)

[http://bioexploratorium.pl/wiki/
Bazy_Danych_i_USlugi_Sieciowe_-_2011z](http://bioexploratorium.pl/wiki/Bazy_Danych_i_USlugi_Sieciowe_-_2011z)