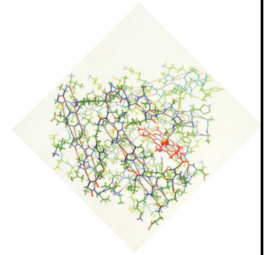




INŻYNIERIA BIAŁEK TERAPEUTYCZNYCH

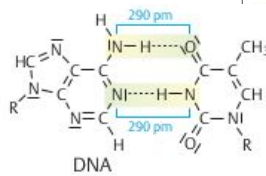
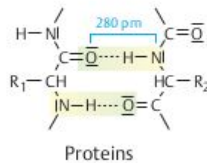
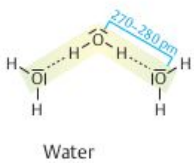
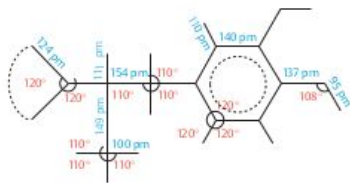


ANDRZEJ ŁYSKOWSKI, DR INŻ.

alyskowski@prz.edu.pl, H.237

Used with permission from the Howard Hughes Medical Institute (www.hmi.org). All rights reserved.

Budowa molekularna



A. Molecule illustrations
Chiral center

1. Formula illustration

2. Ball-and-stick model

B. Bond lengths and angles

C. Bond polarity

1. Partial charges in L-dopa

0.9	2.1	2.5	3.0	3.5	4.0
Na	H	C	N	O	F
1	2	3	4		

2. Electronegativities

1. Donor 2. Acceptor

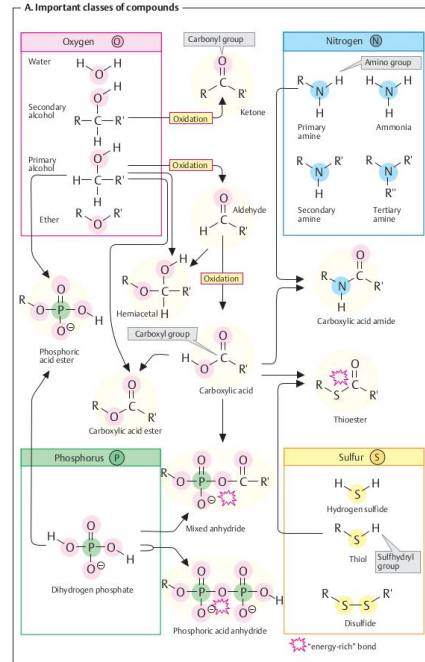
Base Dissociated acid Protonated base

B A — H — B O⁻ H — B⁺

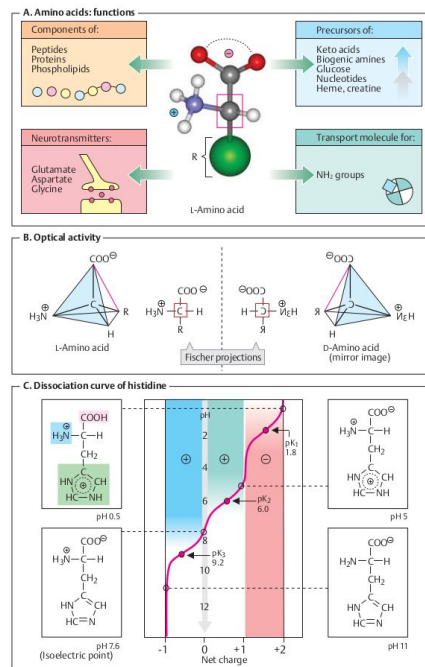
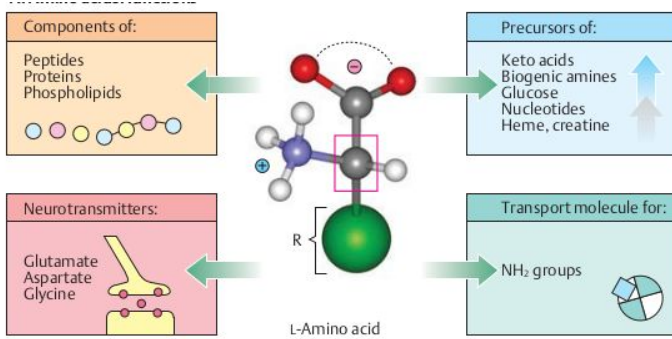
Hydrogen bond Complete reaction

2. Examples

Biocząsteczki i ich właściwości



Aminokwasy



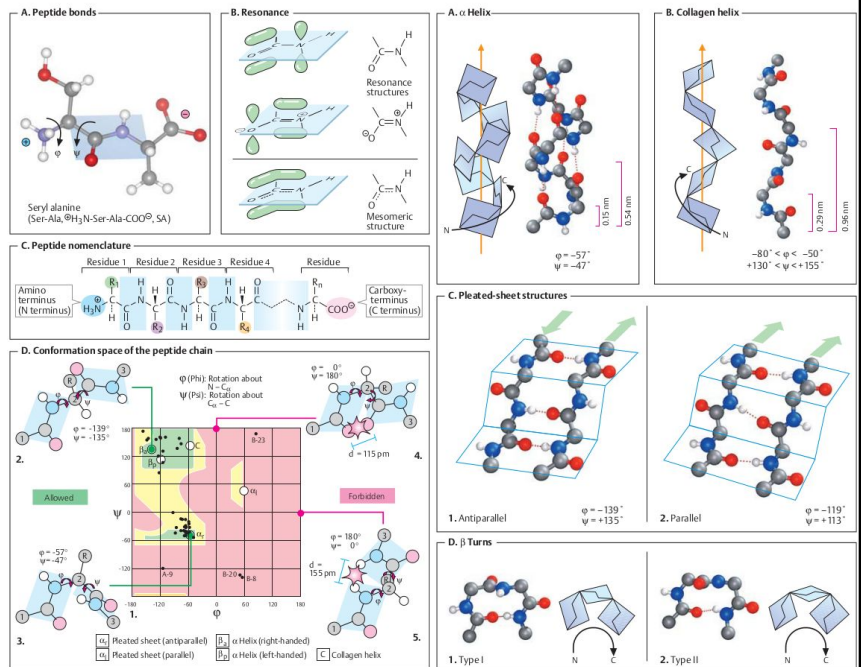
Właściwości aminokwasów

- alifatyczne
- zawierające siarkę
- aromatyczne
- cykliczne
- neutralne
- kwaśne
- zasadowe

A. The proteinogenic amino acids

Aliphatic					Sulfur-containing	
Glycine (Gly, G)	Alanine (Ala, A)	Valine (Val, V)	Leucine (Leu, L)	Isoleucine (Ile, I)	Cysteine (Cys, C)	Methionine (Met, M)
H	CH ₃	H ₃ C-CH CH ₃	CH ₂ H ₃ C-CH CH ₃	H ₃ C-CH-CH CH ₂ CH ₃	CH ₂ SH 8.3	CH ₂ CH ₂ S CH ₃
Polarity						
-2.4	-1.9	-2.0	-2.3	-2.2	-1.2	-1.5
Aromatic			Cyclic	Neutral		
Phenylalanine (Phe, F)	Tyrosine (Tyr, Y)	Tryptophan (Trp, W)	Proline (Pro, P)	Serine (Ser, S)	Threonine (Thr, T)	
+0.8	+6.1	+5.9	+6.0	+5.1	+4.9	
* Essential amino acids □ Chiral center						
Neutral		Acidic		Basic		
Asparagine (Asn, N)	Glutamine (Gln, Q)	Aspartic acid (Asp, D)	Glutamic acid (Glu, E)	Histidine (His, H)	Lysine (Lys, K)	Arginine (Arg, R)
+9.7	+9.4	+11.0	+10.2	+10.3	+15.0	+20.0

Właściwości białek



Struktura białek

A. α-Keratin

Right-handed α-helix
Left-handed superhelix
Protofilament
Intermediary filament
3 nm
10 nm

B. Collagen

1. Triple helix (section)

Gly	Arg	Hyp
Gly	Gln	Arg
Gly	Pro	Hyp
Gly	Pro	Gln
Gly	Ala	Arg

2. Typical sequence
gly X Y

3. Triple helix (view from above)

A. Conformation-stabilizing interactions

Disulfide bond
Polar surface
Apolar core
Metal complex
Hydrogen bond

B. Disulfide bonds

S[C@@H](N)C(=O)O + S[C@@H](N)C(=O)O >> S[C@@H](N)C(=O)O-S-C[C@@H](N)C(=O)O

1. Triple helix (section)

C. Protein dynamics

Mobile domain
Substrate
Adenylate kinase

D. Folding patterns

1. Myoglobin
2. Estrogen receptor (domain)
3. Flavodoxin
4. Transducin (β subunit)

Aktywność biologiczna

A. Structure of insulin

1. Primary structure

2. Secondary and tertiary structure

3. Quaternary structure

B. Insulin (monomer)

1. A-chain, B-chain
2. Invariant residue
3. Apolar side chain, Polar side chain, Involved in subunit interactions

UniProtKB statistics

UniProt BLAST Align Peptide search ID mapping SPARQL



Introduction

This is release 2024_01 of UniProtKB, published on Wed Jan 24 2024.

Previous release statistics are available from the UniProt FTP server.

Throughout this document, whenever a statistic has a corresponding query, a link has been provided. In some instances, due to the nature of the sta

Total number of entries in this release of UniProtKB

Section	Number of entries in total	Number of entries with an annotation update	Number of entries with a sequence update
UniProtKB	250,322,721	161,118,372	1,287
Reviewed (Swiss-Prot)	570,830	433,476	169
Unreviewed (TrEMBL)	249,751,891	160,684,896	1,118

Total number of new entries in this release of UniProtKB

Section	Number of new entries	Number of new sequences
UniProtKB	3,922,146	3,922,117
Reviewed (Swiss-Prot)	415	386
Unreviewed (TrEMBL)	3,921,731	3,921,731

Proteins
UniProt Knowledgebase

Reviewed (Swiss-Prot)
270,830

Unreviewed (TrEMBL)
249,751,891

Protein sets for species with sequenced genomes from across the tree of life

Clusters of protein sequences at 100%, 90% & 50% identity

Non-redundant archive of publicly available protein sequences seen across different databases

Supporting Data

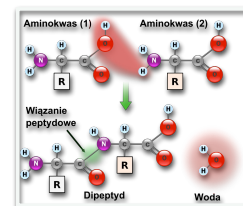
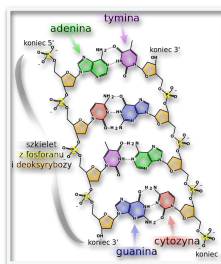
Human diseases | Taxonomy | Keywords | Literature Citations

Cross-referenced databases | Subcellular locations | Automatic annotations: UniRule & ARBA

Dopasowanie sekwencji

Klasyfikacja przyrównań (ang. sequence alignment):

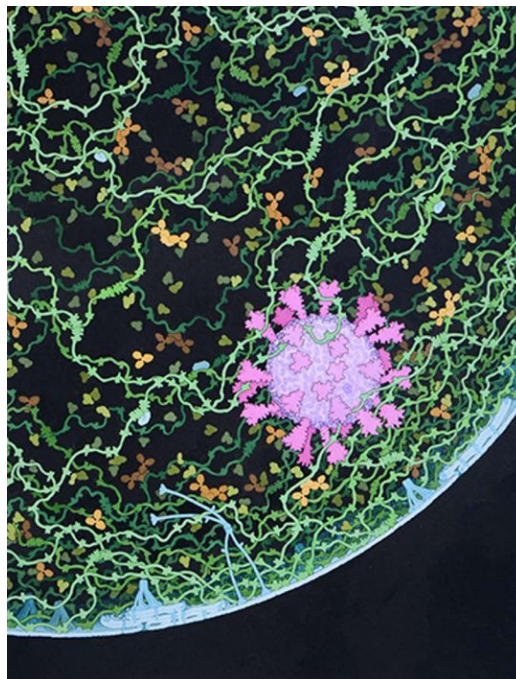
- nukleotydydowej: cztery elementy porównawcze w tripletach.
- sekwencji aminokwasowej: 20 jednostek porównawczych.
- strukturze przestrzennej: struktura I, II, III, IV rzędowa definiująca przestrzenne ułożenie atomów.



Dane strukturalne

Protein Data Bank (1971, 7 struktur)

- Pierwsze źródło danych biologicznych w historii
- Globalne archiwum struktur:
 - białek
 - DNA/RNA
- Otwarty dostęp do >210 000 struktur
- Kordynatorzy:
 - RCSB PDB (US)
 - PRBe (EMBL-EBI)
 - PDBj (Japonia)
 - PDBc (Chiny)
 -
 - EMDB (3DEM)
 - BMRB (NMR)

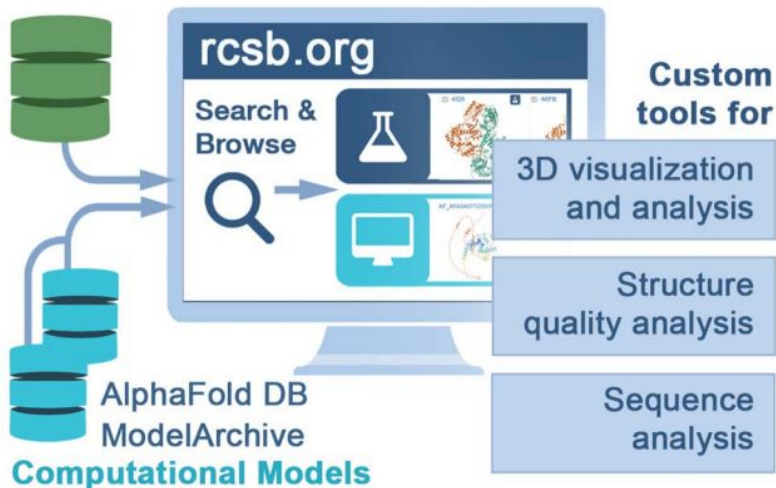


Dostępne dane

rscsb.org:

- >210 000 struktur PDB
- >1 000 000 modeli komputerowych (Computed Structure Models, CSM)
 - AlphaFold DB
 - ModelArchive
- Narzędzia do:
 - wizualizacji i analizy
 - oceny jakości struktur
 - analizy sekwencji

Experimental Models Protein Data Bank



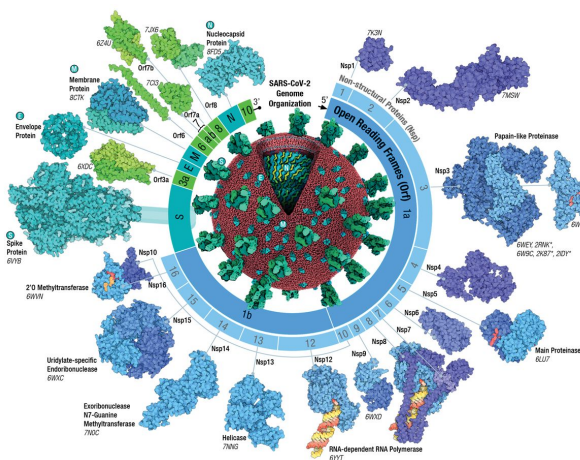
rscsb.org



Znaczenie danych strukturalnych

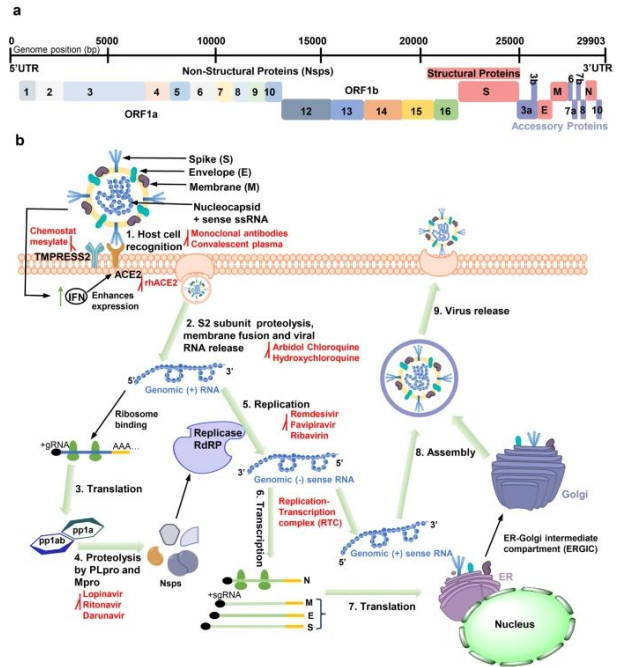
- SARS-CoV 2002
>170 SARS-CoV structuresPD
- MERS-CoV 2012
>100 MERS-CoV structuresP
- COVID-19 2019
>3,000 SARS-CoV-2 structures

- Określenie pokrewieństwa
- Projektowanie szczepionek
- Identyfikacja leków przeciwwirusowych



Koronawirus SARS-CoV-2 genom

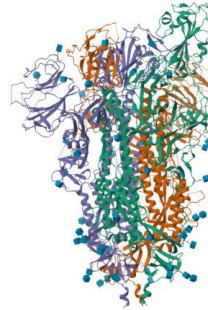
- Genom: RNA
- Białka strukturalne kodowane oddzielnie
- Białka funkcjonalne kodowane jako polipeptyd
- MP: main protease
- PL: papain like proteinase



Yan, W., Zheng, Y., Zeng, X. et al. Structural biology of SARS-CoV-2: open the door for novel therapies. *Sig Transduct Target Ther* 7, 26 (2022). <https://doi.org/10.1038/s41392-022-00884-5>

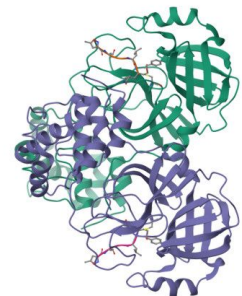
Strategie wykorzystujące dane strukturalne

- Szczepionka
- identyfikacja immunogenego elementu
- Zastosowanie leków przeciwwirusowych

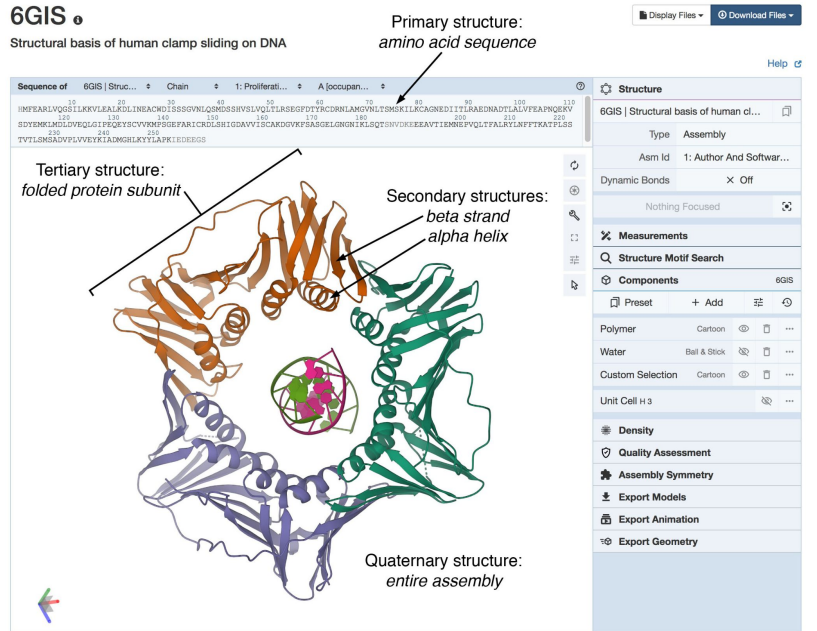


- 6VSB Prefusion 2019-nCoV spike glycoprotein with a single receptor-binding domain up

- 6LU7 The crystal structure of COVID-19 main protease in complex with an inhibitor N3



Organizacja molekularna

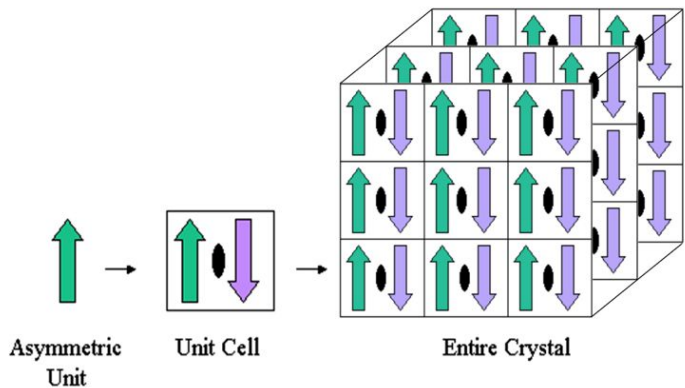


Organizacja molekularna

Jednostka asymetryczna jest najmniejszą częścią struktury krystalicznej, do której można zastosować operacje symetrii w celu wygenerowania kompletnej komórki elementarnej (jednostki powtarzalnej kryształu).

Operacje symetrii najczęściej spotykane w kryształach makrocząsteczek biologicznych to rotacje, translacje i osie śrubowe (kombinacje rotacji i translacji).

Zastosowanie operacji symetrii krystalograficznej do jednostki asymetrycznej daje jedną komórkę elementarną, która po przetłumaczeniu w trzech wymiarach tworzy cały kryształ.



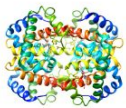

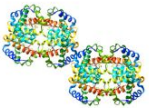
Znaczenie biologiczne

Krystaliczna jednostka asymetryczna może zawierać:

- jeden zespół biologiczny
- część zespołu biologicznego
- wiele zespołów biologicznych

Zawartość jednostki asymetrycznej zależy od pozycji krystalizowanej cząsteczki i jej konformacji w komórce elementarnej. W zależności od warunków krystalizacji i lokalnego upakowania mogą wystąpić dwa różne scenariusze:

- Kopie makrocząsteczki lub kompleksu w krystalicznej komórce elementarnej mają identyczne konformacje i zajmują pozycje związane z symetrią. W rezultacie zespół biologiczny może składać się z jednej kopii makrocząsteczki/kompleksu lub może składać się z dwóch lub więcej cząsteczek/kompleksów powiązanych symetrią, tworzących większy zespół.
- Kopie makrocząsteczki lub kompleksu przyjmują nieco inne konformacje i zajmują unikalne pozycje w krystalicznej jednostce asymetrycznej. W rezultacie każda z różnych pozycji makrocząsteczki/kompleksu może odpowiadać strukturalnie podobnym, ale nie identycznym zespołom biologicznym.


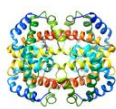
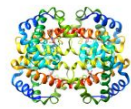
Asymmetric unit with one biological assembly	Asymmetric unit with a portion of a biological assembly	Asymmetric unit with multiple biological assemblies
		
Entry 2hhb contains one hemoglobin molecule (4 chains) in the asymmetric unit.	Entry 1out contains half a hemoglobin molecule (2 chains) in the asymmetric unit. A crystallographic two-fold axis generates the other 2 chains of the hemoglobin molecule.	Entry 1hv4 contains two hemoglobin molecules (8 chains) in the asymmetric unit.

Interpretacja komórki elementarnej

Zespół biologiczny (czasami określane również jako jednostka biologiczna) to zespół makromolekularny, który okazał się być lub uważa się, że jest funkcjonalną formą cząsteczki. Na przykład, funkcjonalna forma hemoglobiny ma cztery łańcuchy.

W zależności od konkretnej struktury krystalicznej, operacje symetrii składające się z rotacji, translacji lub ich kombinacji mogą wymagać wykonania w celu uzyskania kompletnego zespołu biologicznego. Alternatywnie, może być konieczne wybranie podzbioru zdeponowanych współrzędnych w celu przedstawienia zespołu biologicznego. W ten sposób zespół biologiczny może być zbudowany z:


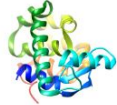

- jednej kopii jednostki asymetrycznej
- wielu kopii jednostki asymetrycznej
- części jednostki asymetrycznej

Biological assembly composed of one copy of the asymmetric unit	Biological assembly composed of multiple copies of the asymmetric unit	Multiple biological assemblies in the asymmetric unit
		
In entry 2hhb , the biological assembly is equivalent to the asymmetric unit.	In entry 1out the biological assembly includes two asymmetric units.	In entry 1hv4 the biological assembly is one-half of the asymmetric unit.
No operations are necessary.	Application of a crystallographic symmetry operation (a 180 rotation around a crystallographic two-fold axis) produces the complete biological assembly.	The entry contains two structurally similar, but not entirely identical copies of the biological assembly within the crystal asymmetric unit.

(Nad)interpretacja komórki elementarnej

Cząsteczka może czasami wydawać się multimeryczna w kryształach na podstawie upakowania kryształów. Jednak może nie być żadnych dowodów lub biologicznego znaczenia na poparcie stanu multimerycznego w roztworze.

Gdy wpis jest przetwarzany, wszystkie prawdopodobne złożenia są obliczane na podstawie zakopanej powierzchni i energii interakcji. Te przewidywane złożenia mogą, ale nie muszą pokrywać się z tym, co autor uważa za biologicznie istotne dla cząsteczki. Zespoły biologiczne zgłoszone we wpisie zawierają uwagę wyjaśniającą, czy jest to "dostarczone przez autora", "określone przez oprogramowanie", czy oba.

Asymmetric unit (monomer)	Author & Software Determined Biological Assembly (monomer)	Software Determined Biological Assembly (dimer)
		
The asymmetric unit is a monomer. These are the deposited coordinates.	The "author provided" and "software determined" biological assemblies are both monomers.	The software, PISA, predicts that this molecule may also form a dimer. Hence the second biological assembly is only "software determined".

Identyfikacja elementów kluczowych

- Struktury białek zorganizowane są w motywy
- CATH
Class(C), Architecture(A),
Topology(T), and Homologous
Superfamily(H)
- SCOP
Structural Classification of Proteins



CATH

- Wyszukiwanie
 - Text or ID
 - FASTA
 - Structure
- Wynik
 - "domeny"
 - "elementy funkcjonalne"
 - "ścieżka ewolucyjna"

CATH / Gene3D v4.3
151 million protein domains classified into 5,841 superfamilies

Search by keywords, PDB code, GO term, etc.

CATH annotations for the 21 model organisms predicted by AlphaFold (v2) are available to download (doi:10.1038/s42003-023-04488-9). Core classification files for the latest version of CATH-Plus (v4.3) are available to download. Daily updates of our very latest classifications are also available.

3D Structure
Find out what 3D structure your protein adopts.

Protein Evolution
Learn about a particular protein family and how it evolved.

Protein Function
Investigate the function of your protein and how it evolved.

Conserved Sites
Look at protein sites that are highly conserved and implicated in function.

Download Data
Download data files and query CATH via web services.

Learn more
Find out how CATH is created and maintained, how to link to CATH and more.

What is CATH-Gene3D?
CATH is a classification of protein structures downloaded from the Protein Data Bank. We group protein domains into superfamilies when there is sufficient evidence they have diverged from a common ancestor.

Latest Release Statistics

	CATH-Plus 4.3.0	CATH (stable snapshots)
PDB Release	01-07-2019	5 months ago
Domains	580235	536613
Superfamilies	5401	6031
Annotated PDBs	150005	158379

Gene3D v21
Protein Sequences: 32,005,354
CATH Domain Predictions: 151,913,797

6VSI
>6VSB
glycopr
syndror

CATH Domain
Home / Superfamily 3.30.70

CATH Superfamily 3.30.70.1840
Spike protein, C-terminal core receptor binding subdomain

DOMAIN LINKS
Summary
Structure
Sequence
Neighbourhood

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.

- Positive spike proteins
- Positive spike proteins
- Both glycoprotein

Superfamily Summary
A general summary of information for this superfamily.

Structures
Domains: 58
Domain clusters (>95% seq ID): 6
Domain clusters (>35% seq ID): 2
Unique PDBs: 28

Alignments
Structural Clusters (SA): 1
Structural Clusters (SA): 1
FunFam Clusters: 3

Function
Unique EC: 10
Unique GO: 10

Taxonomy
Unique Species: 121

GO Diversity
Unique GO annotations: 10
Unique GO terms: 10

EC Diversity
Unique EC annotations: 1
Unique EC terms: 1

Species Diversity
Unique species annotations: 1
Unique species: 10

Sequence/Structure Diversity
Overview of the sequence / structure diversity of this superfamily compared to other superfamilies in CATH. Click on the chart to view the data in more detail.

Structure Diversity vs Sequence Diversity chart showing 3.30.70.1840 highlighted.

Search Results
1 keywords
cture
JD8
PDBSum
Proteopedia
RAY DIFFRACTION
scherichia
API Help
Structure of Severe Acute Respiratory Syndrome Coronavirus Receptor-binding Domain Complexed with Neutralizing Antibody
abakaran, P., Gan, J., Jing, Y., Zhu, Z., Choudhry, , Xiao, X., Ji, X., Dimitrov, S.
3iol.Chem.

Functional families

Browse Functional Families

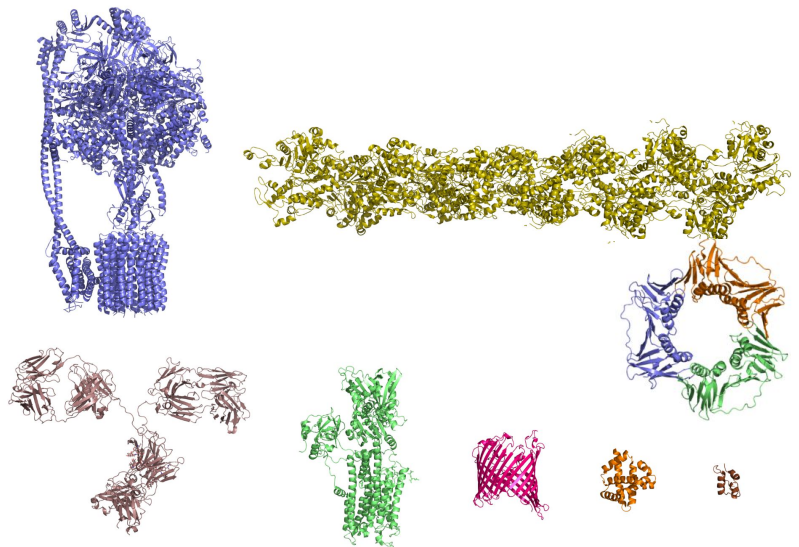
Displaying 2 FunFams

Filter FunFams by Keywords

ID	Function Family (FunFam) Name	Total Sequences	Enzyme?	Structure?	Structural Representative	PDB Sites?	Alignment Diversity (0-100)
1	Spike glycoprotein	32			-	-	5.8
2	Spike glycoprotein	19		3D	3r4dD00	-	65.4

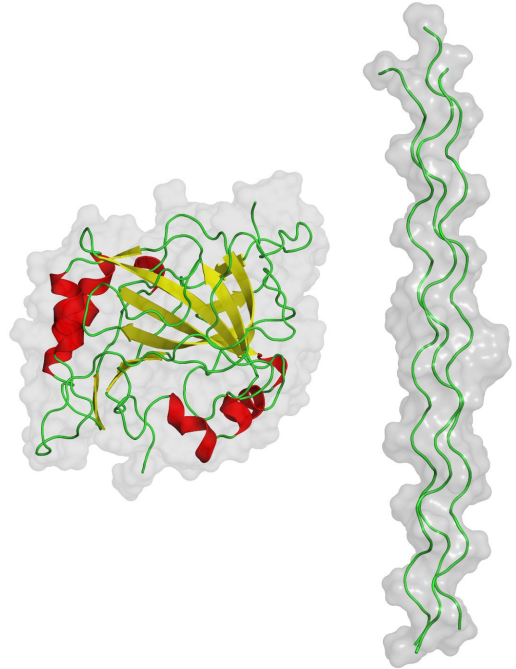
Różnorodność białek

- ATP synthase, 4814 aa
- Actin filament, 4512 aa
- Proliferating cell nuclear antigen, 754 aa
- Insulin, 21 aa
- Myoglobin, 153 aa
- Porin, 413 aa
- Ca²⁺ transporter, 1000 aa
- Antibody, 1316 aa



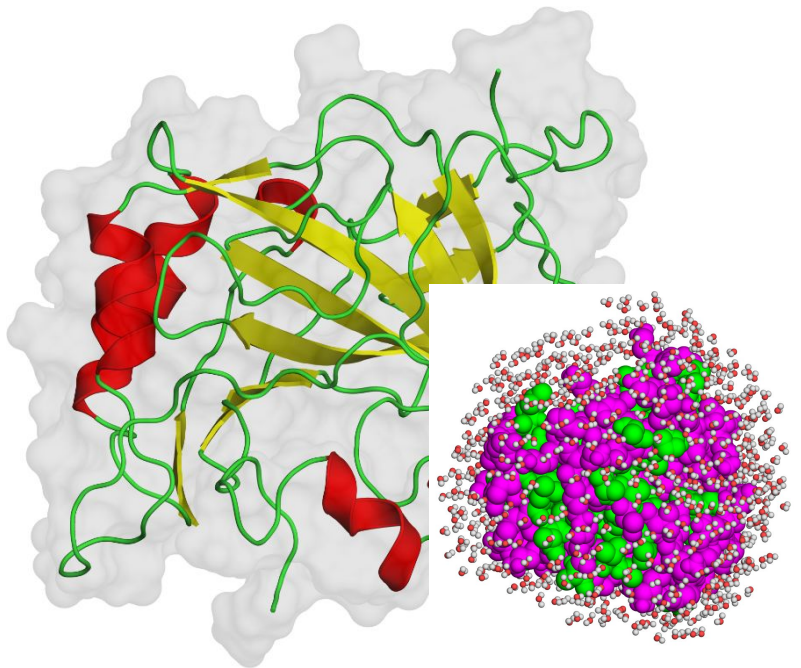
Strukturalne rodzaje białek

- Globularne
białko globularne, białko o kształcie kulistym lub sferoidalnym, zbudowane z gęsto połączonych lub pozwijanych łańcuchów peptydowych; wartość stosunku osiowego cząsteczki nie przekracza 3 : 1 lub 4 : 1; do b. g. należy ogromna większość białek, m.in. albuminy i globuliny oraz wszystkie enzymy.
- Fibylarne
białka proste o strukturze włóknkowej stanowiące podstawowy materiał budulcowy organizmów zwierzęcych. Są to typowe białka o budowie włóknistej, dzięki temu pełnią funkcje podporowe. Do tej grupy białek należy keratyna, kolagen, miozyna i fibroina.



Białka globularne

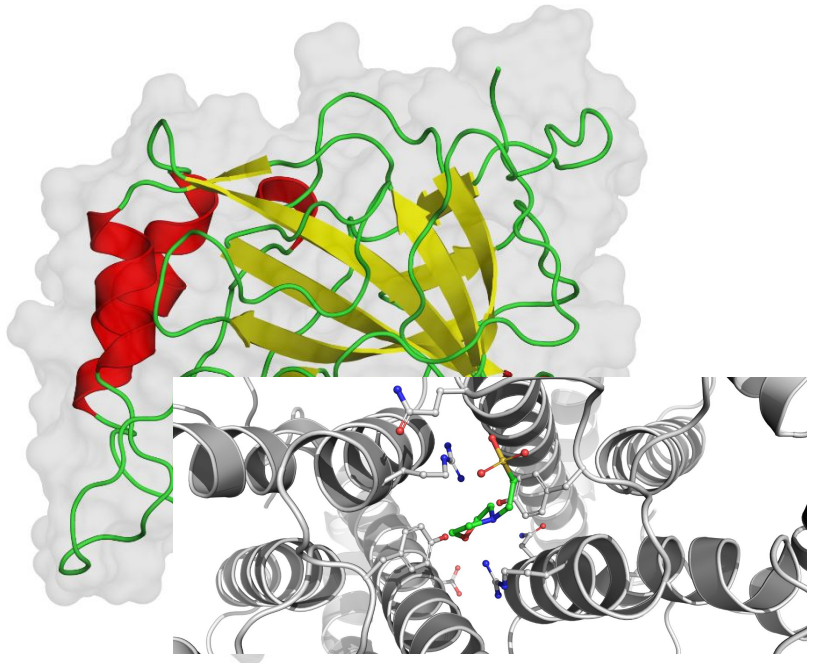
- Kompaktowe (lokalizacja cytoplazmatyczna)
- Rdzeń hydrofobowy (zapewnia stabilność)
- Powierzchnia hydrofilowa (zapewnia rozpuszczalność)



Właściwości

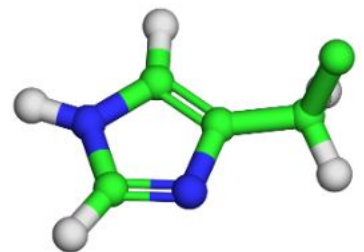
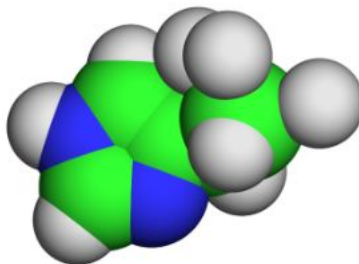
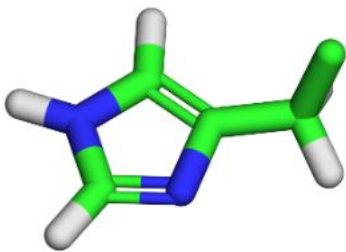
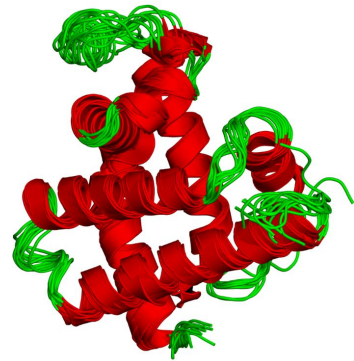
Faldowanie białka jest niezbędne dla:

- formacji przestrzennych podstaw aktywności biologicznej:
 - miejsca wiązania
 - miejsca katalitycznego
 - miejsca regulacji/aktywacji
- wiąże się z:
 - zdolnością katalityczną
 - funkcjami transportowymi
 - 'odpornością'
 - regulacją aktywności

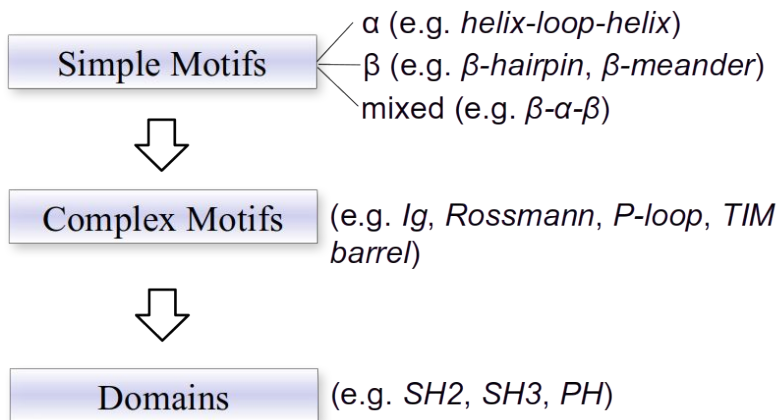


Stabilność

- Interakcje niekowalencyjne
- Kowalencyjne wiązania S-S (mostki disiarczkowe, cysteine)
- Elastyczność w obszarach aktywności biologicznej



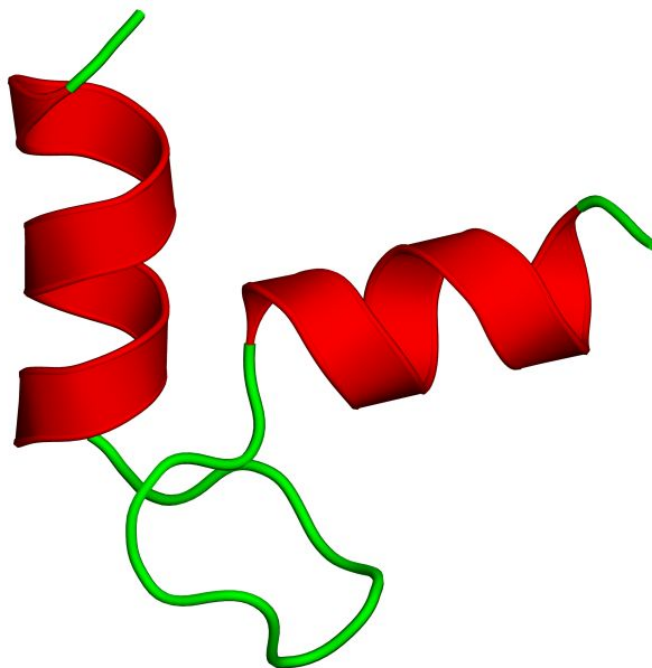
Architektura
białek globularnych



HLH
helix-loop-helix

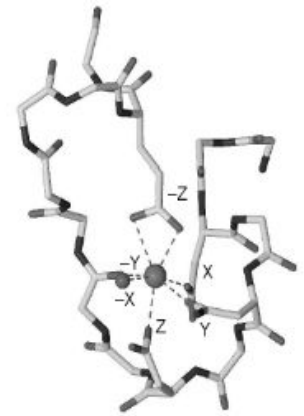
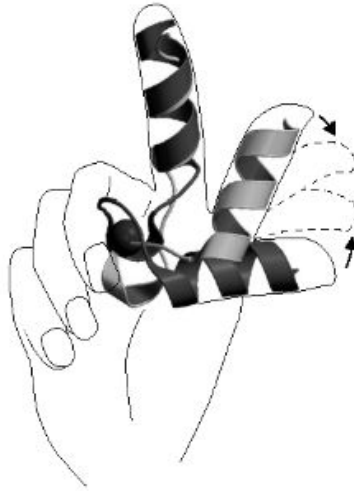
EF-hand (wiązanie Ca^{2+})

bHLH (wiązanie DNA)



EF-hand (wiązanie Ca^{2+})

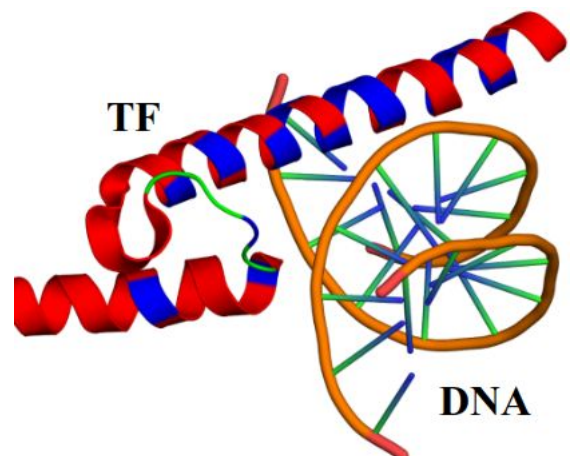
- Ca^{2+} wiąże się do polarnych atomów w obszarze pętli i cząsteczek wody
- Wiązanie wymaga zmiany położenia helisy



Curr. Opin. Struct. Biol. (2000) 10: 637-643

bHLH

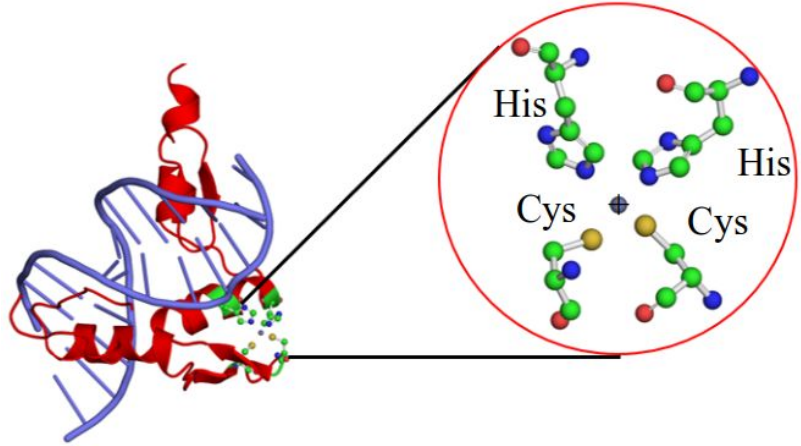
- Występuje w czynnikach transkrypcyjnych (transcription factors, TF)
- Rozpoznaje specyficzne sekwencje DNA
- Aminokwasy 'miejsca aktywnego': zasadowe



HLH

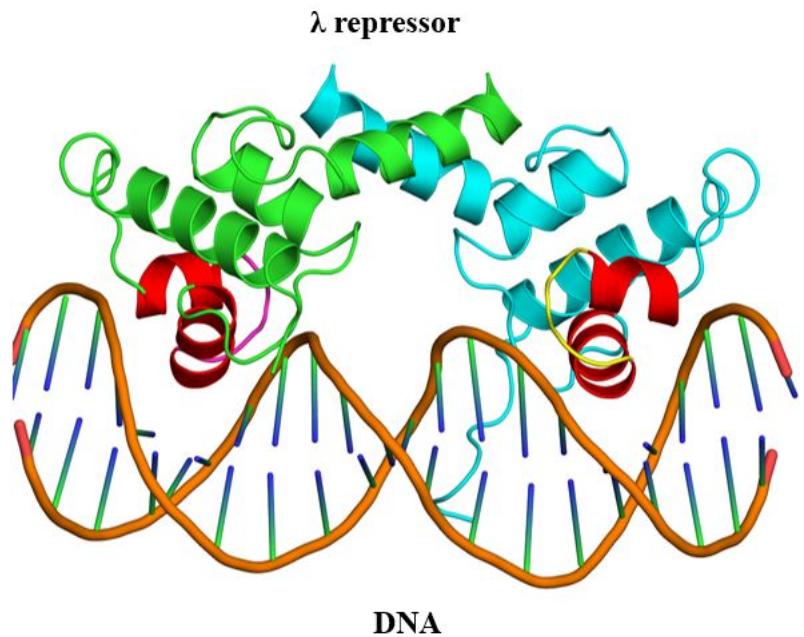
palec cynkowy

- Palec cynkowy, motyw palca cynkowego (ang. zinc finger domain) – rodzaj domeny białkowej występującej w białkach wiążących DNA i biorący bezpośredni udział w związaniu cząsteczki kwasu nukleinowego przez białko. Palec cynkowy składa się z dwóch antyrównoległych β -katek i α -helisy. Obecność jonu cynku (Zn^{2+}) jest kluczowa dla stabilności domeny – w przypadku jego braku nie dochodzi do powstania wystarczająco dużej domeny posiadającej hydrofobowy rdzeń i przez to funkcjonalnej.



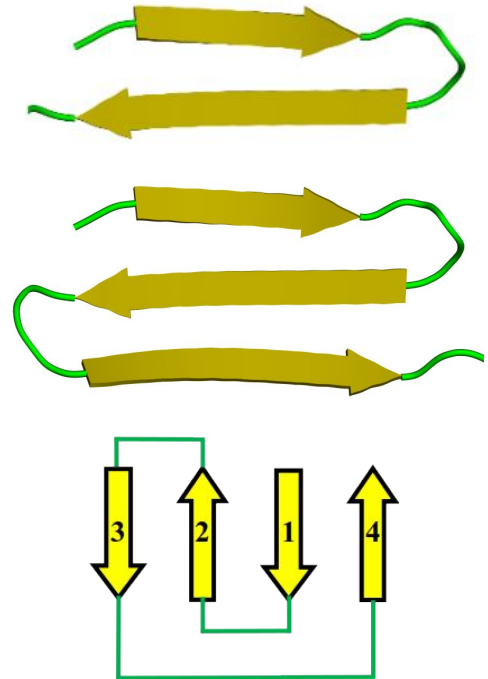
HTH

- Helisa-Zwrot-Helisa (HTH, ang. Helix-Turn-Helix) wiąże się do głównego rowka DNA



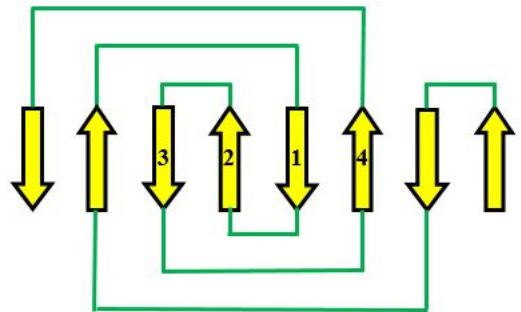
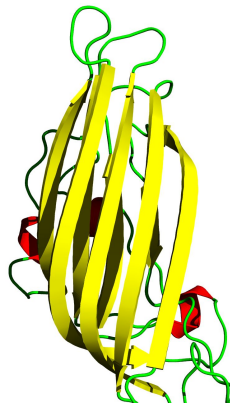
Motywy β

- β spinka
- β meander
- motyw klucza greckiego



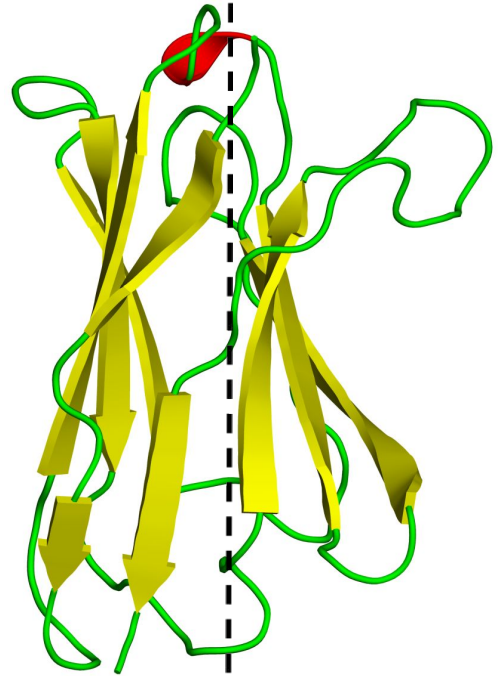
Motywy β

- Jelly roll
Jelly roll lub Swiss roll fold to układ białkowy składająca się z ośmiu nici beta ułożonych w dwa czteroniciowe arkusze. Nazwa tej struktury została wprowadzona przez Jane S. Richardson w 1981 roku, odzwierciedlając jej podobieństwo do galaretki lub szwajcarskiego ciasta. Fałd jest rozwinięciem greckiego motywu klucza i jest czasami uważany za formę beczki beta. Jest on bardzo powszechny w białkach wirusowych, w szczególności w białkach kapsydu wirusowego. Łącznie, struktury galaretki i klucza greckiego stanowią około 30% wszystkich białek beta z adnotacjami w bazie danych Structural Classification of Proteins (SCOP).



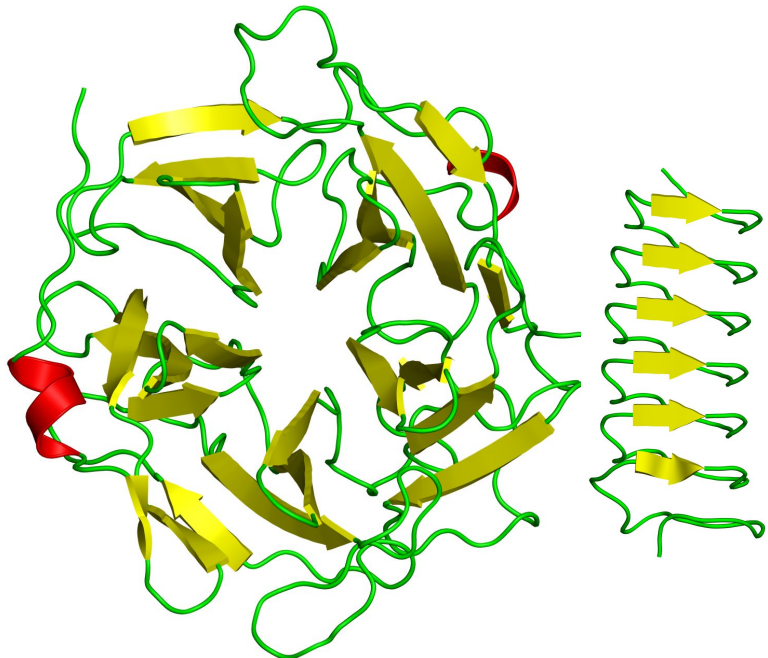
Motywy β

- Domeny beta-sandwich, β -sandwich składające się z 80 do 350 aminokwasów występują powszechnie w białkach. Charakteryzują się one dwoma przeciwnymi antyrównoległymi arkuszami beta (β -sheets). Liczba pasm występujących w takich domenach może się różnić w zależności od białka. Domeny β -sandwich są podzielone na wiele różnych układów. Typ immunoglobulinowy występujący w przeciwciałach składa się z kanapkowego układu 7 i 9 antyrównoległych pasm β ułożonych w dwóch arkuszach β o topologii klucza greckiego.



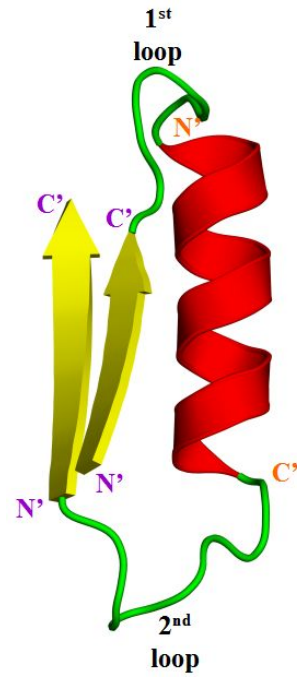
Motywy β

- śmigło β beta-propeller (β -propeller) jest typem architektury białek all- β , charakteryzującym się 4 do 8 wysoce symetrycznymi arkuszami beta w kształcie ostrza, ułożonymi toroidalnie wokół centralnej osi. Razem arkusze beta tworzą lejkowate miejsce aktywne.
- β helisa to tandemowa struktura powtórzenia białka utworzona przez połączenie równoległych arkuszy beta w spiralny wzór z dwoma lub trzema ścianami. Helisa beta jest rodzajem solenoidowej domeny białkowej. Struktura ta jest stabilizowana przez międzywęzłowe wiązania wodorowe, interakcje białko-białko, a czasami związane jony metali. Zidentyfikowano zarówno lewo-, jak i prawoskrętne helisy beta.



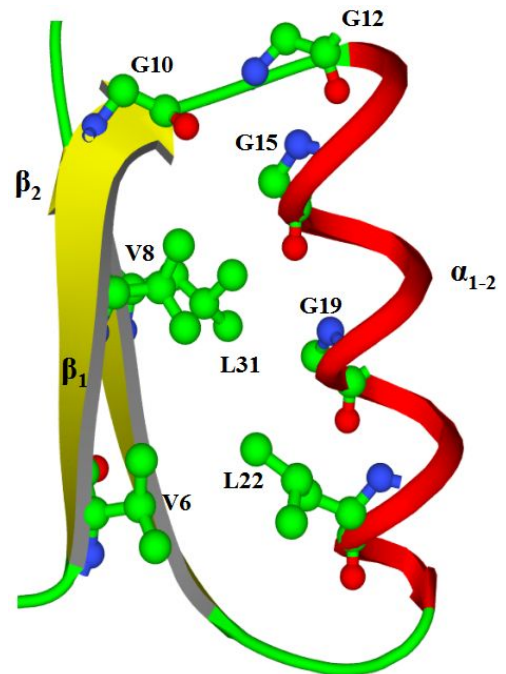
Motywy mieszane

- β - α - β
 - niezwykle popularny(!)
 - pozwala ukrywać reszty aminokwasowe
 - charakterystyczne miejsce wiązania



Stabilność

- Interakcje helisa-harmonijka efektywnie 'pakują' łańcuch boczne

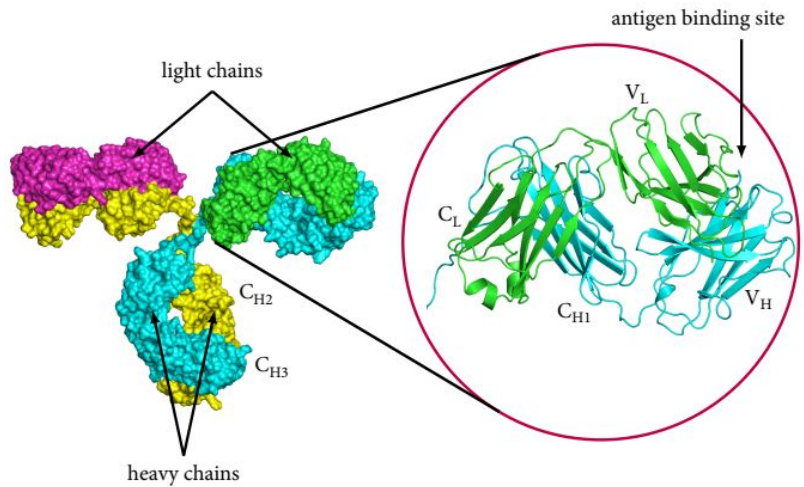


Domena immunoglobulinowa

Domena immunoglobulinowa, znana również jako motyw immunoglobulinowy, jest rodzajem domeny białkowej, która składa się z dwuwarstwowej kanapki 7-9 antyrównoległych pasm β ułożonych w dwa arkusze β o topologii klucza greckiego składającej się z około 125 aminokwasów.

Szkielet wielokrotnie przełącza się między dwoma arkuszami β . Zazwyczaj wzór jest następujący (N-końcowa β -spinka do włosów w arkuszu 1)-(β -spinka do włosów w arkuszu 2)-(β -nić w arkuszu 1)-(C-końcowa β -spinka do włosów w arkuszu 2). Przejścia między arkuszami tworzą "X", tak że N- i C-końcowe spinki do włosów są skierowane do siebie.

Członkowie nadrodziny immunoglobulin występują w setkach białek o różnych funkcjach. Przykłady obejmują przeciwciała, gigantyczną kinazę mięśniową titinę i receptorowe kinazy tyrozynowe. Domeny podobne do immunoglobulin mogą być zaangażowane w interakcje białko-białko i białko-ligand.

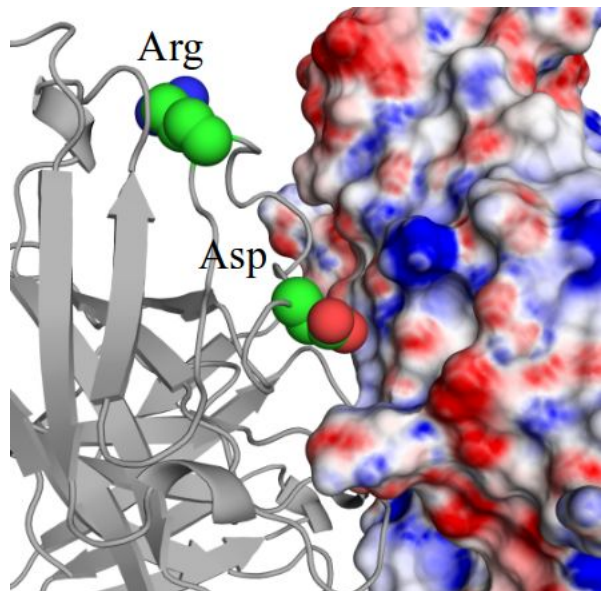


Fałd immunoglobulinowy

Splot immunoglobulinowy, fałd immunoglobulinowy lub domena immunoglobulinowa (domena Ig) – struktura białkowa charakterystyczna dla cząsteczek z nadrodziny immunoglobulin, składającą się z dwóch równoległych harmonijek beta. Wyróżnia się 4 typy domen immunoglobulinowych:

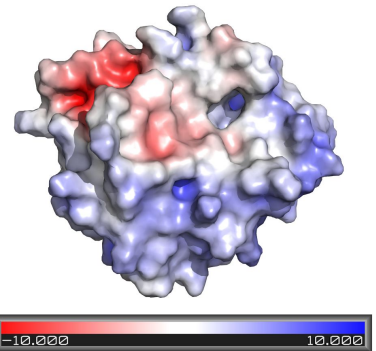
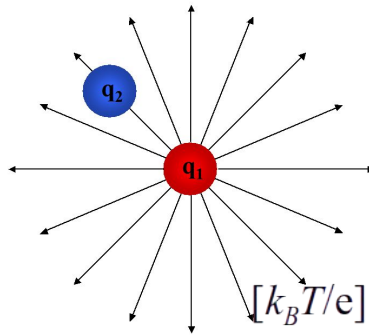
- domeny IgV (znane także jako domeny V lub domeny części zmiennych) – mają największe rozmiary, a typowym przykładem są domeny części zmiennych przeciwciał. Jedna z beta-harmonijek beta ma 4, a druga 5 pasm.
- domeny IgC (znane także jako domeny C lub domeny części stałych) – mają mniejsze rozmiary, niż domeny V i budowę charakterystyczną dla domen Ig znajdujących się w częściach stałych przeciwciał. W obydwu harmonijkach beta znajdują się po 4 pasma aminokwasowe. Takie klasyczne domeny IgC nazywa się domenami IgC1. Do domen IgC zalicza się też domeny o charakterze IgV, ale rozmiarach porównywalnych z klasycznymi domenami IgC1. Noszą one nazwę domen IgC2.
- domeny IgI (z ang. intermediate) – mają cechy nie pozwalające zaklasyfikować ich do żadnej z powyższych grup[3].

Całość, niezależnie od rodzaju domeny, jest stabilizowana przez mostki dwusiarczkowe, a w skład struktury wchodzi 60-100 aminokwasów.



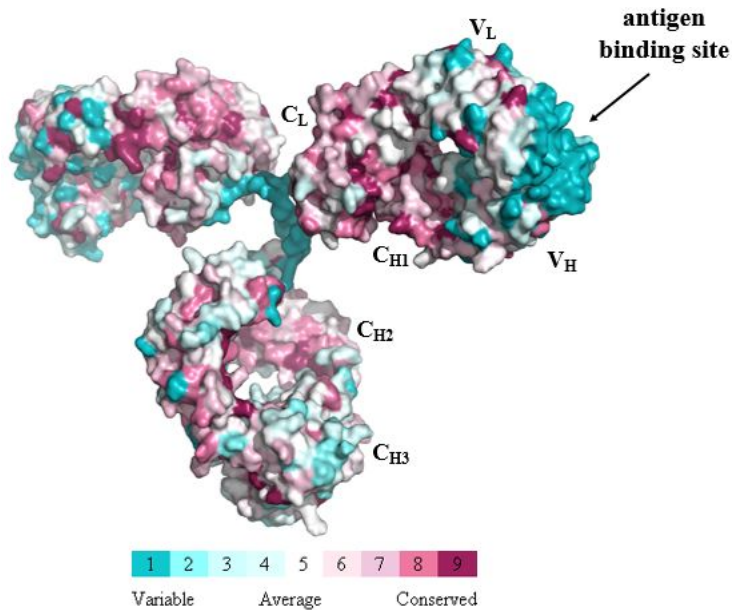
Potencjał elektrostatyczny

- Potencjał elektrostatyczny białek spowodowany naładowanymi łańcuchami bocznymi i związanymi jonami - odgrywa rolę np. w fałdowaniu i stabilności białek, katalizie enzymatycznej lub specyficznym rozpoznawaniu białek.



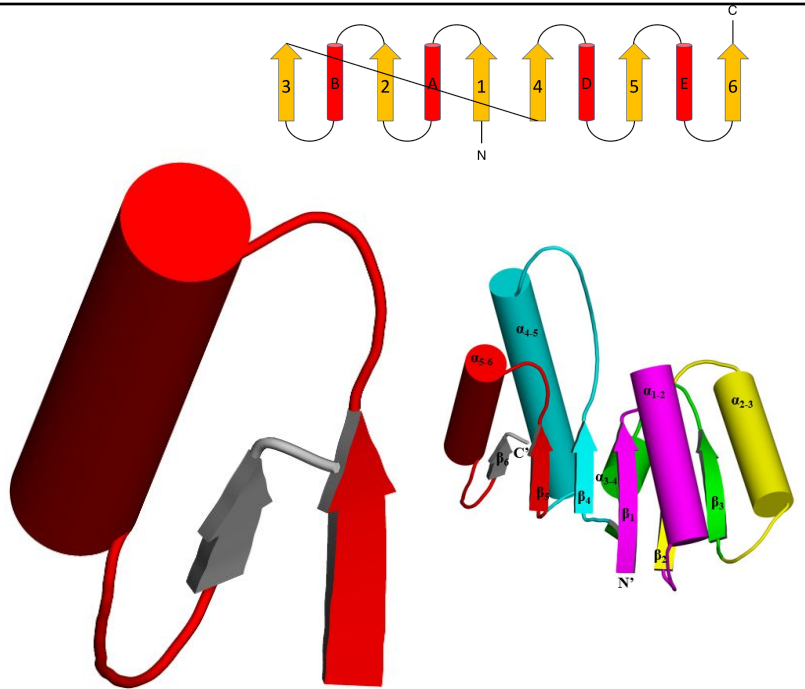
Zmienność ewolucyjna

- Motyw - TAK
- Miejsce wiązania - NIE



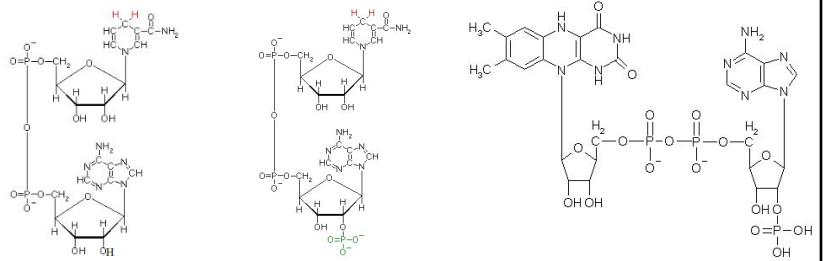
Motyw Rossmann

- Fald Rossmanna to trzeciorzędowy fald występujący w białkach wiążących nukleotydy, takich jak kofaktory enzymatyczne FAD, NAD⁺ i NADP⁺. Fald ten składa się z naprzemiennych nici beta i segmentów helikalnych alfa, w których nici beta są ze sobą związane wodorem, tworząc wydłużony arkusz beta, a helisy alfa otaczają obie powierzchnie arkusza, tworząc trójwarstwową kanapkę.
- Klasyczny fald Rossmanna zawiera sześć nici beta, podczas gdy falda podobne do faldy Rossmanna, czasami określane jako falda Rossmanna, zawierają tylko pięć nici.
- Białka faldowe Rossmanna i Rossmannoida są niezwykle powszechne. Stanowią one 20% białek o znanych strukturach w Protein Data Bank i występują w ponad 38% szlaków metabolicznych KEGG. Fald ten jest niezwykle wszechstronny, ponieważ może pomieścić szeroki zakres ligandów.
- Mogą one funkcjonować jako enzymy metaboliczne, białka wiążące DNA/RNA i białka regulatorowe oprócz tradycyjnej roli.



Motyw Rossmann

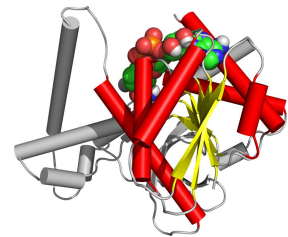
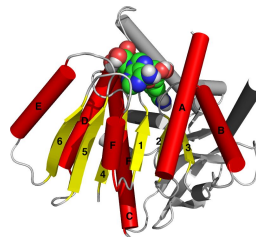
- Kofaktory



NADH

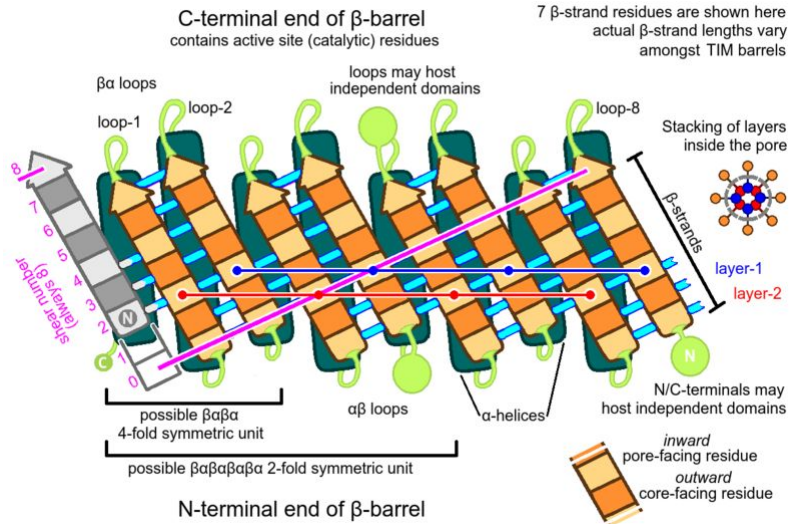
NADPH

FADH₂

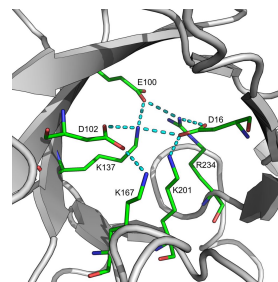
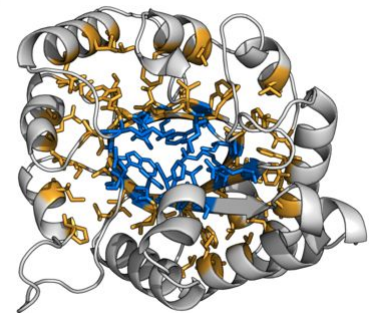
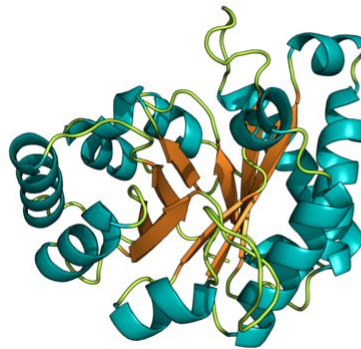
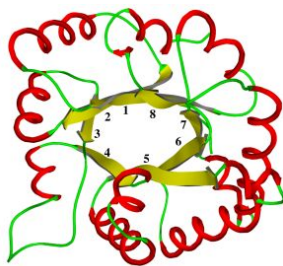


TIM barrel

- Beczka TIM (triose-phosphate isomerase, izomeraza triozowo-fosforanowa), znana również jako baryłka alfa/beta to konserwatywny fałd białkowy składający się z ośmiu helis alfa (helis α) i ośmiu równoległych pasm beta (pasm β), które występują naprzemiennie wzdłuż szkieletu peptydowego.
- Beczki TIM są wszechobecne, a około 10% wszystkich enzymów przyjmuje ten fałd.
- Ponadto pięć z siedmiu klas enzymów komisji enzymatycznej (EC) obejmuje białka baryłkowe TIM.
- Fałd baryłkowy TIM jest ewolucyjnie starożytny, a wielu jego członków ma dziś niewielkie podobieństwo sekwencji.

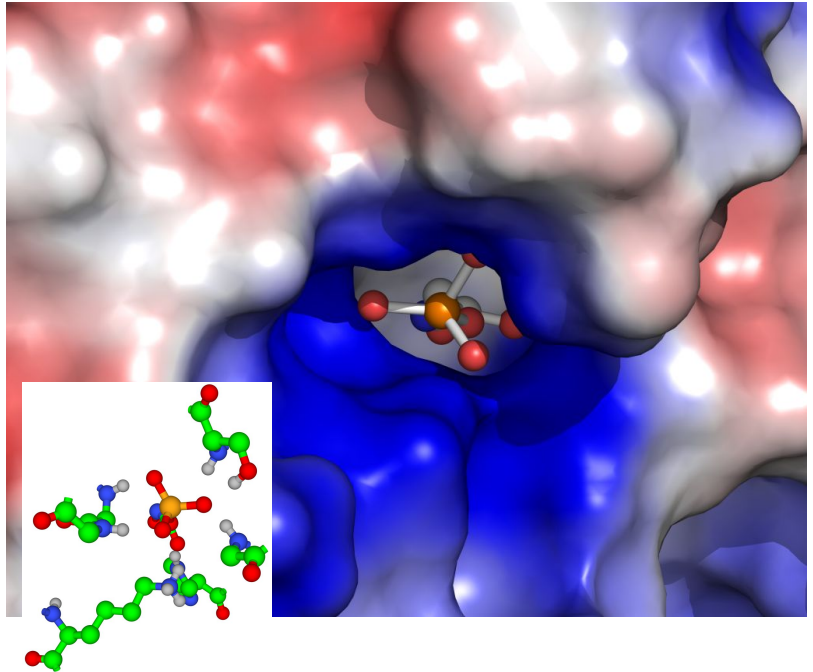


Beczka TIM



Beczka TIM

Miejsce aktywne



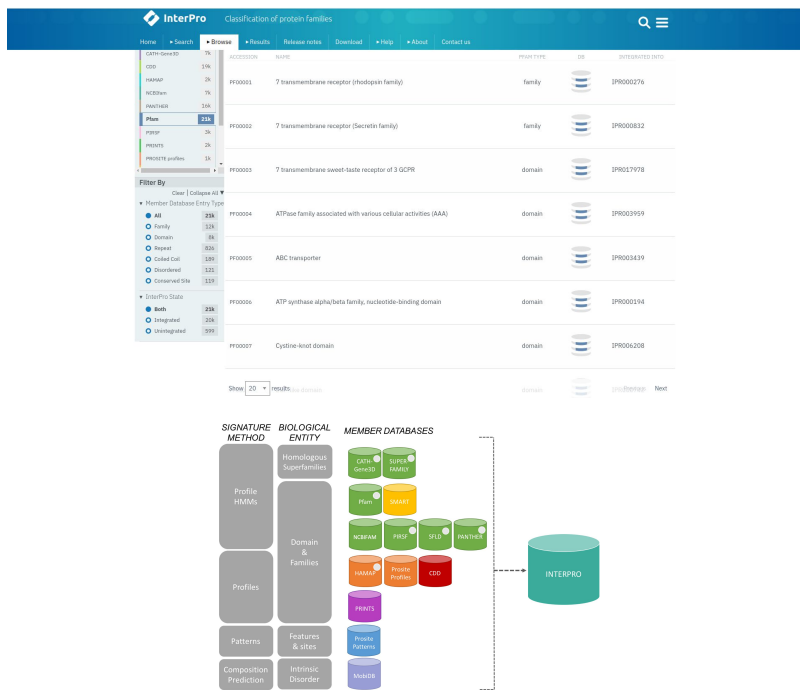
Domena

- Motyw strukturalny (lub funkcjonalny) białek.
- Powtarzalny układ (motyw) aminokwasów zachowany w strukturze lub sekwencji.

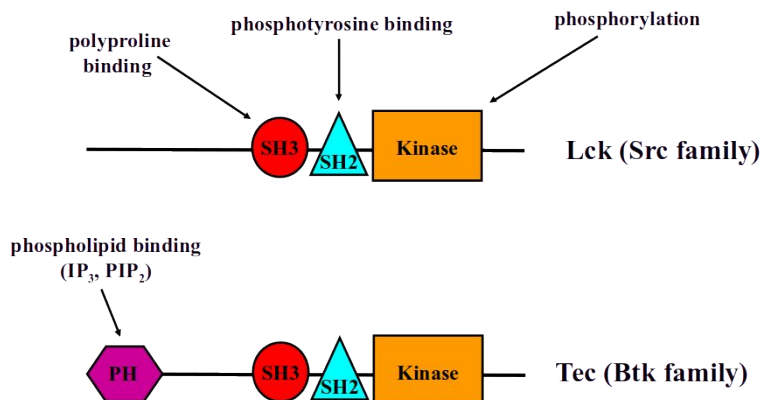
- Funkcja... często nieznaną.
- Występowanie... zazwyczaj 2-5 na białko.

InterPro/Pfam

- InterPro**
 to baza danych rodzin białek, domen białkowych i miejsc funkcjonalnych, w której możliwe do zidentyfikowania cechy znalezione w znanych białkach można zastosować do nowych sekwencji białkowych w celu ich funkcjonalnego scharakteryzowania.
- Pfam**
 to baza danych rodzin białek, która zawiera ich adnotacje i wielokrotne wyrównania sekwencji wygenerowane przy użyciu ukrytych modeli Markowa. Wersja Pfam 36.0, została wydana we wrześniu 2023 roku i zawiera 20 795 rodzin.



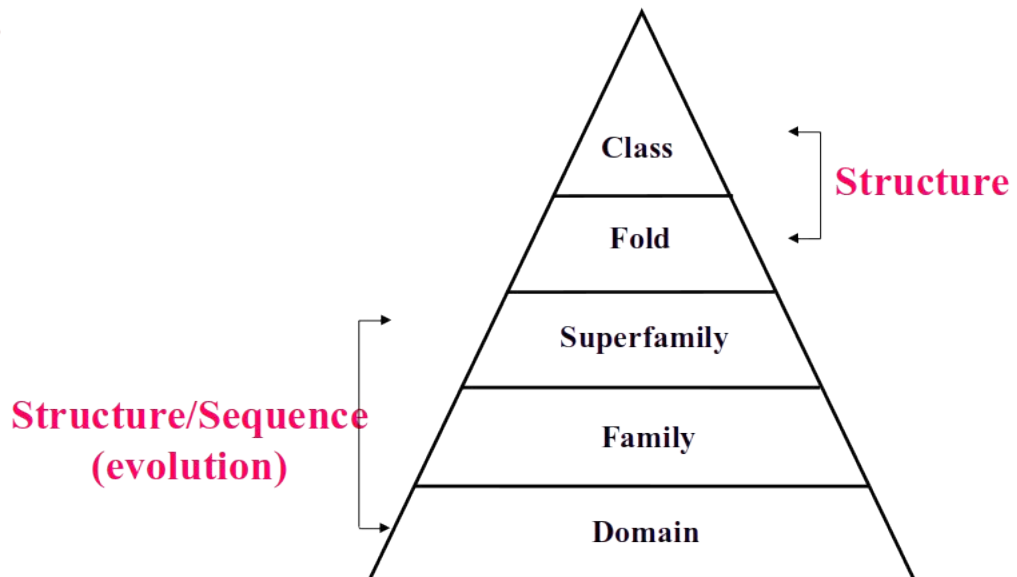
Modularność białek



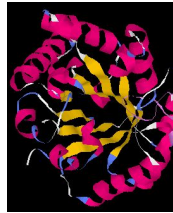
Klasyfikacja pokrewieństwa

- Rodzina
 - $\geq 40\%$ zgodności sekwencji
 - bardzo podobna struktura, pokrewieństwo ewolucyjne
- Super-rodzina
 - $\sim 20\text{-}30\%$ zgodności sekwencji
 - podobna struktura, dalekie pokrewieństwo ewolucyjne
- Klasa
 - podobna organizacja na poziomie sekwencji (domeny)

SCOP



Klasyfikacja

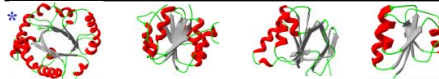


Classes:

1. [All alpha proteins](#) [46456] (226)
2. [All beta proteins](#) [48724] (149)
3. [Alpha and beta proteins \(a/b\)](#) [51349] (134)
Mainly parallel beta sheets (beta-alpha-beta units)
4. [Alpha and beta proteins \(a+b\)](#) [53931] (286)
Mainly antiparallel beta sheets (segregated alpha and beta regions)
5. [Multi-domain proteins \(alpha and beta\)](#) [56572] (48)
Folds consisting of two or more domains belonging to different classes
6. [Membrane and cell surface proteins and peptides](#) [56835] (49)
Does not include proteins in the immune system
7. [Small proteins](#) [56992] (79)
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] (24)
Not a true class
10. [Peptides](#) [58231] (116)
Peptides and fragments. Not a true class
11. [Designed proteins](#) [58788] (42)
Experimental structures of proteins with essentially non-natural sequences

Kalsyfikacja

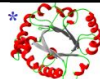
Class: α/β proteins



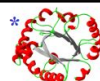
Fold: TIM β/α-barrels



Superfamily: TIM



Family: TIM

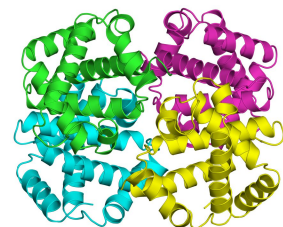
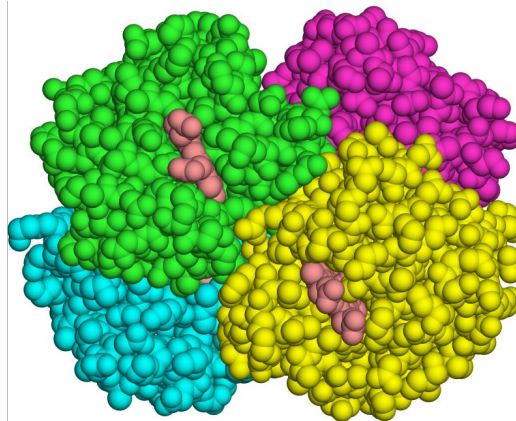


Aspekty inżynieryjne

- Mutacja
 - miejsce aktywne:
zmiana aktywności +/-
 - miejsce wiązania:
zmiana powinowactwa +/-
 - struktura:
zmian właściwości makroskopowych +/-

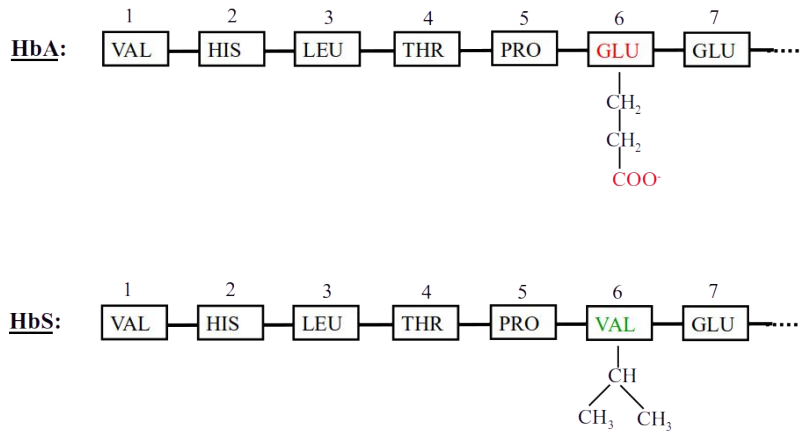
Niedokrwistość sierpowata

- Niedokrwistość sierpowata, anemia sierpowata (łac. anaemia drepanocytica, ang. sickle cell anemia) rodzaj wrodzonej niedokrwistości spowodowanej nieprawidłową budową hemoglobiny.

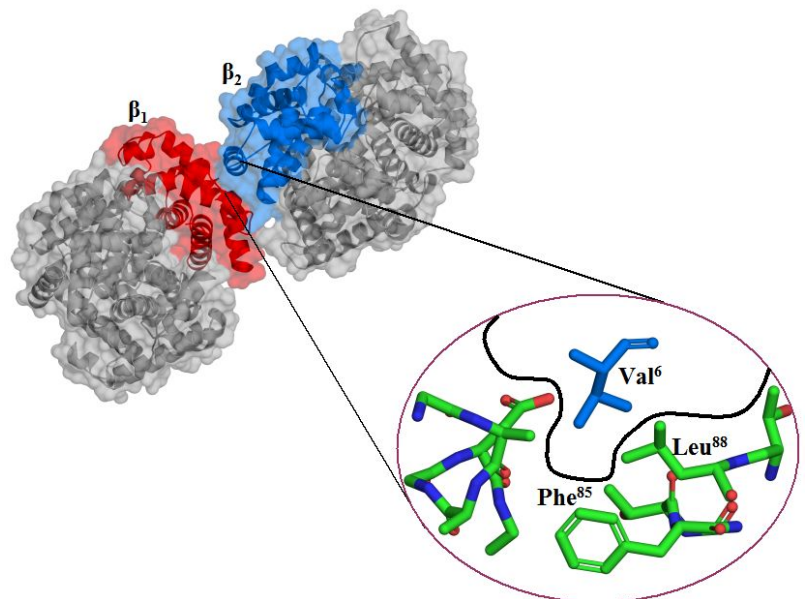


Niedokrwiistość sierpowata

- Mutacja punktowa w genie łańcucha β (HBB) hemoglobiny powoduje zmianę pojedynczego aminokwasu w sekwencji białka (z kwasu glutaminowego na walinę, w pozycji 6 od końca NH₂).
- Hemoglobinę z tak zmienioną, nieprawidłową strukturą 1-rzędową określa się jako hemoglobinę S (HbS) w przeciwieństwie do normalnej, występującej u dorosłych hemoglobiny A (HbA). Hemoglobina S charakteryzuje się zmienionymi w porównaniu z hemoglobiną A własnościami fizykochemicznymi.



Niedokrwiistość sierpowata



Aspekty inżynieryjne

Co zachować?

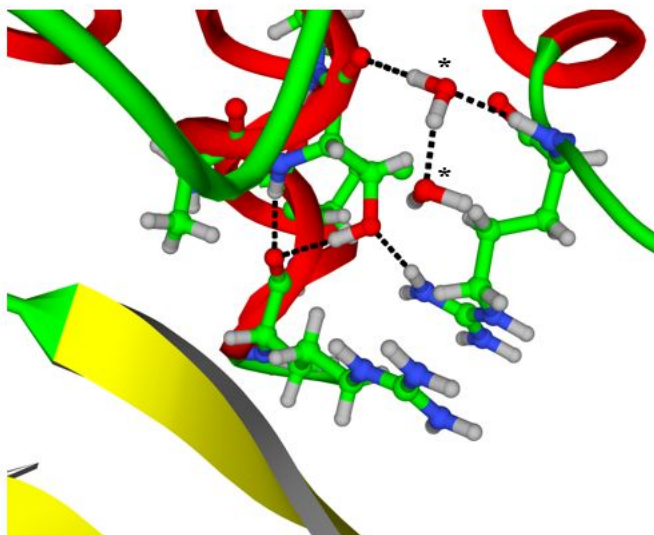
- Obszary funkcjonalne (centrum aktywne, etc.)
- Rdzeń białka
- Integralność struktury drugorzędowej

Jak mutować?

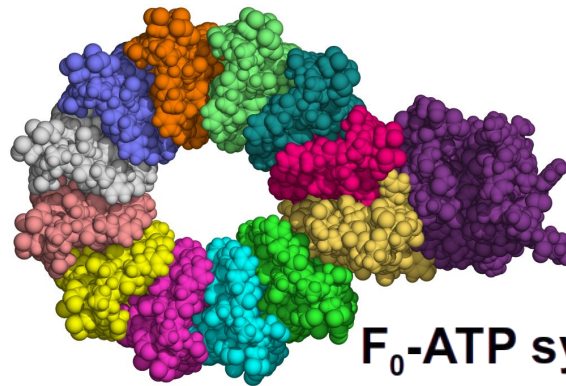
- Konserwatywnie wykorzystując pokrewieństwo aminokwasów.
- Liberalnie kompensując wprowadzane zmiany.

Woda

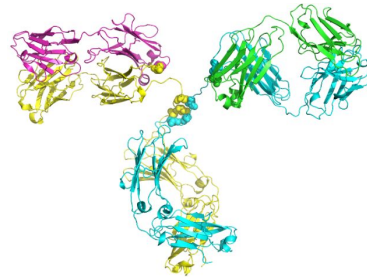
- Maskuje grupy polarne
- Pomaga w wiązaniu związków małowczątkowych
- Pomaga w katalizie (kwas/zasada, polaryzacja wiązań)
- Pomaga w transporcie



Struktura czwartorzędowa

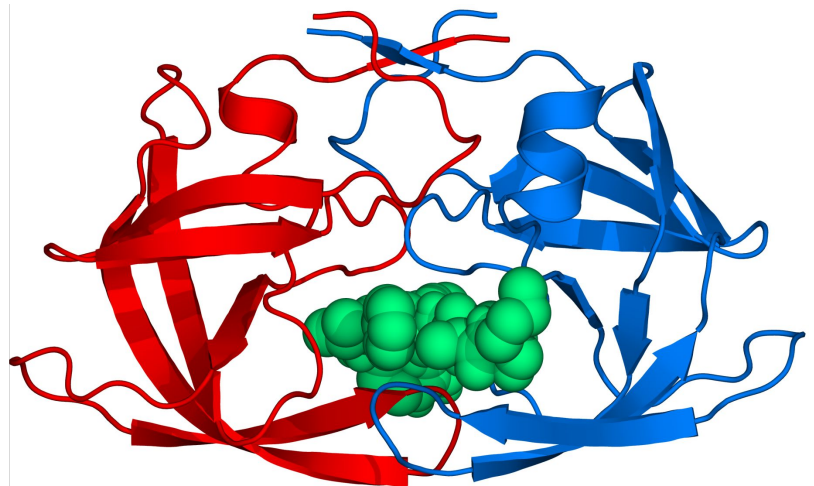


F₀-ATP synthase



Znaczenie
struktury czwartorzędowej

- Formacja miejsca aktywnego
- Poprawa stabilności (np. insulina)



The HIV-1 aspartyl protease

Insulina

