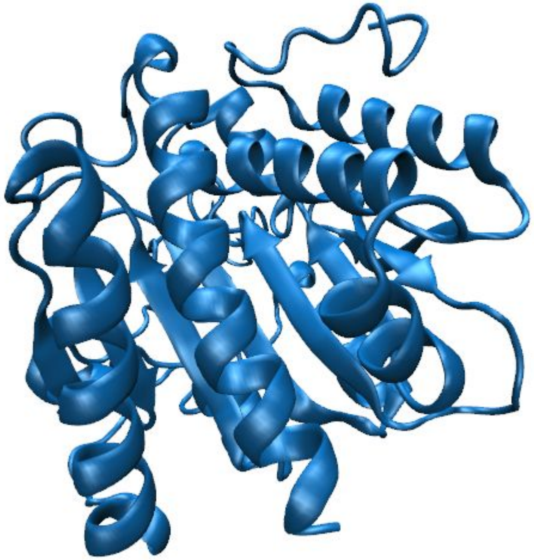
Carboxypeptidases

Carlota Bellot Herrero

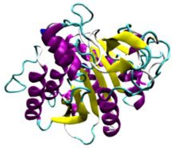
Marcel Lucas Sánchez

Irene Ortega González

Clàudia Prat Gibert



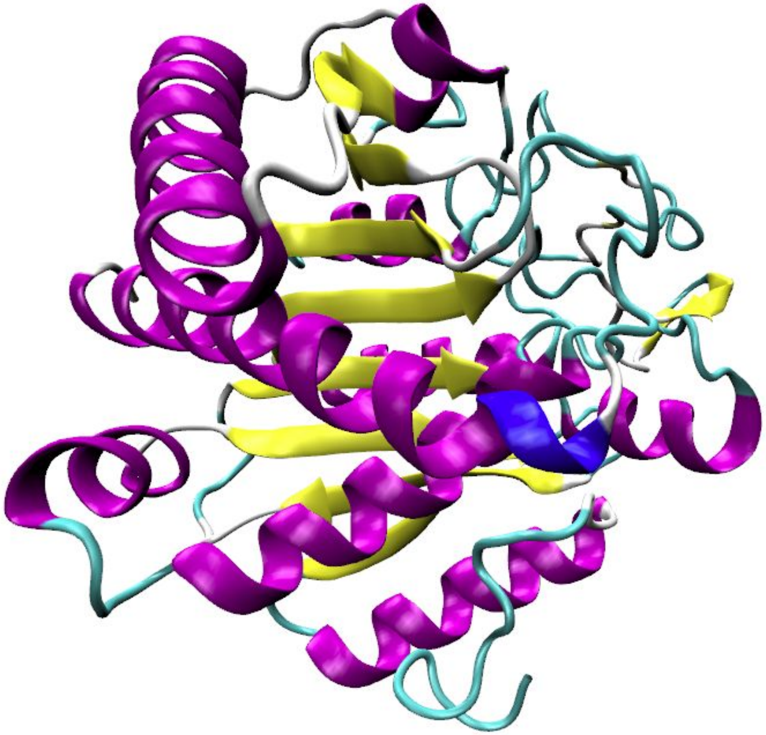
Index



- **Introduction**
  - Definition and function
  - Classification
- **General structure**
  - Procarboxypeptidase
  - Rossman Fold-like
- **Active site**
  - Location
  - Zinc-binding residues
  - Subsites
- **Enzymatic reaction**
  - Substrate-binding interactions
  - Catalytic reaction
- **Human carboxypeptidases**
  - Pancreatic
  - Substrate specificity
  - Pancreatic vs Regulatory
- **Conclusions**

What are carboxypeptidases?

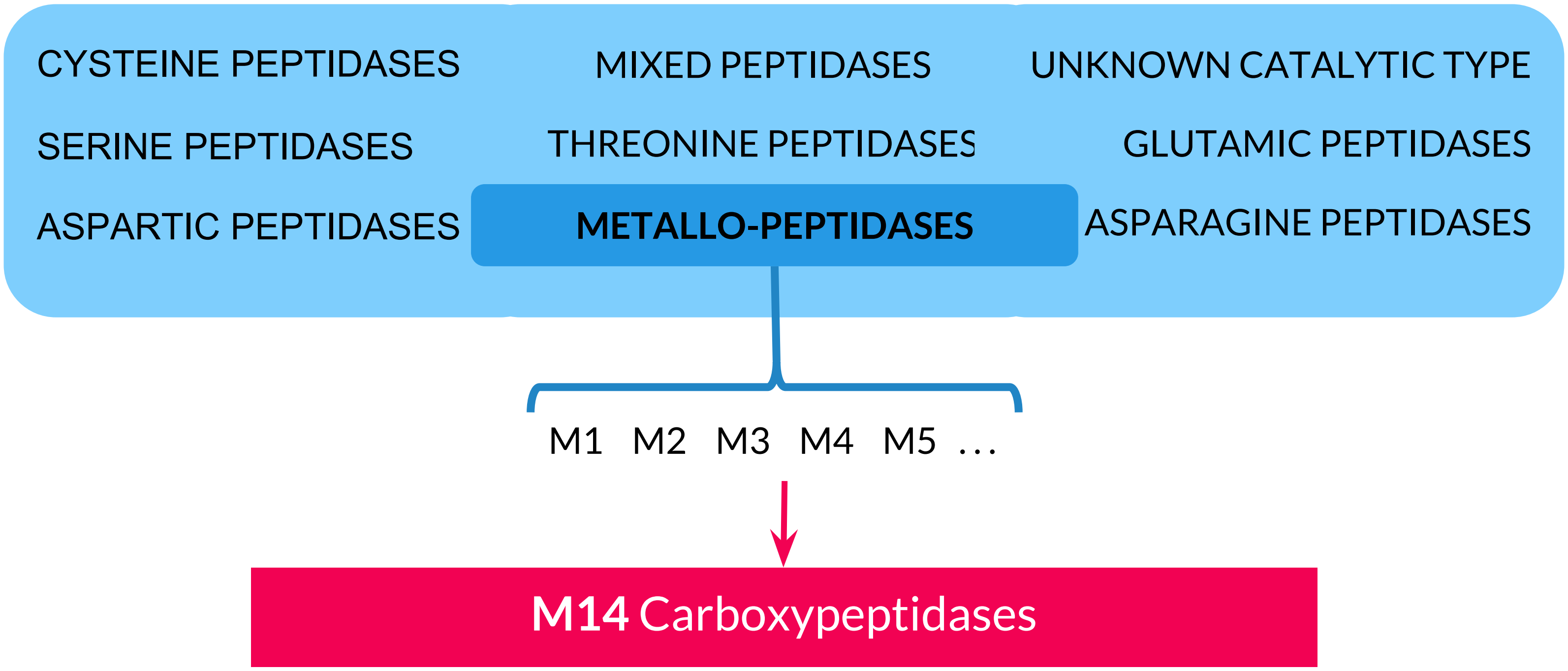
- Exopeptidase
- Cleavage at C-terminus
- Widely distributed



Classification

- CYSTEINE PEPTIDASES
- SERINE PEPTIDASES
- ASPARTIC PEPTIDASES
- MIXED PEPTIDASES
- THREONINE PEPTIDASES
- METALLO-PEPTIDASES
- UNKNOWN CATALYTIC TYPE
- GLUTAMIC PEPTIDASES
- ASPARAGINE PEPTIDASES

Classification



CYSTEINE PEPTIDASES

SERINE PEPTIDASES

ASPARTIC PEPTIDASES

MIXED PEPTIDASES

THREONINE PEPTIDASES

METALLO-PEPTIDASES

UNKNOWN CATALYTIC TYPE

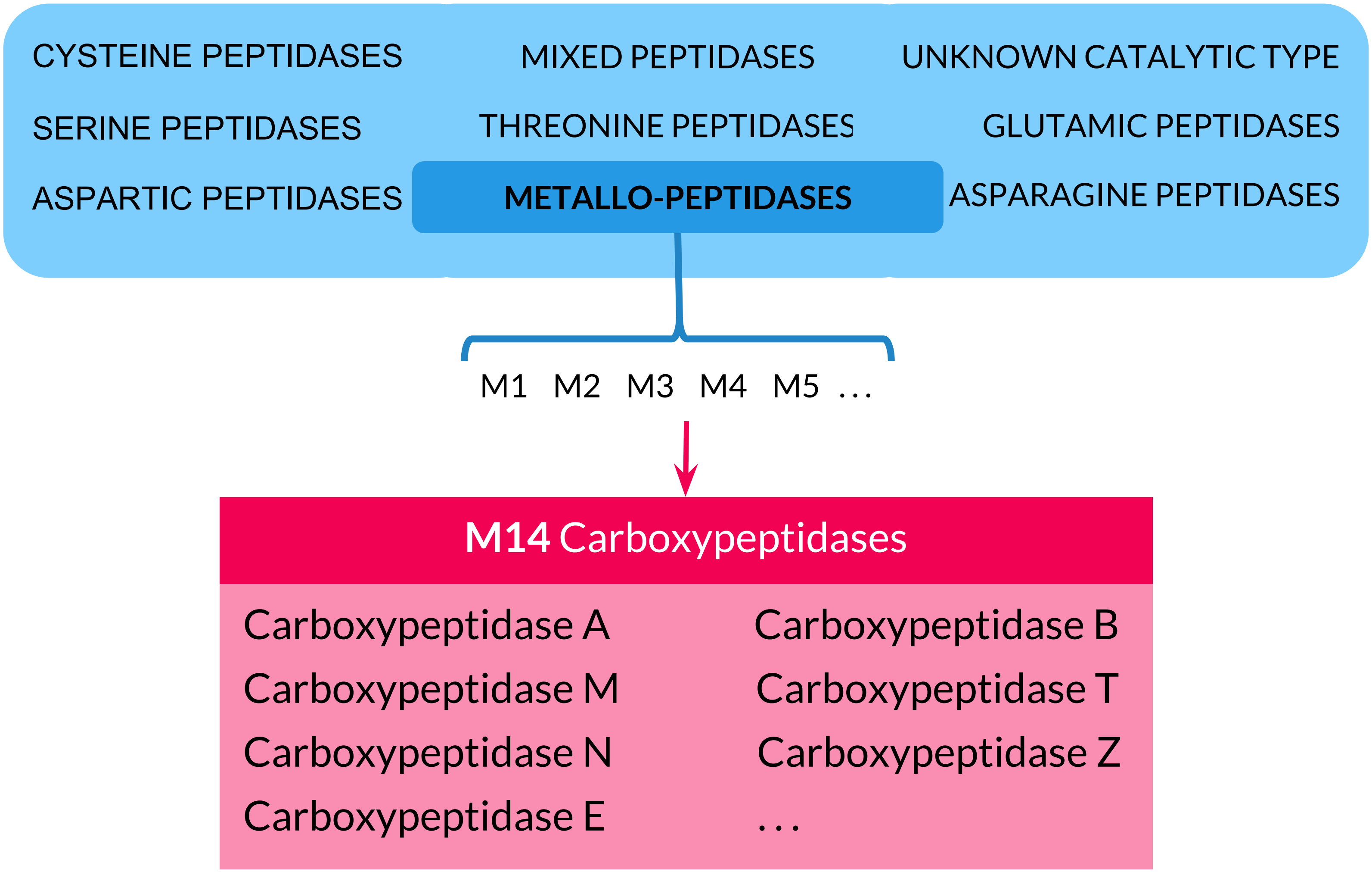
GLUTAMIC PEPTIDASES

ASPARAGINE PEPTIDASES

M1 M2 M3 M4 M5 . . .

M14 Carboxypeptidases

Classification



CYSTEINE PEPTIDASES

MIXED PEPTIDASES

UNKNOWN CATALYTIC TYPE

SERINE PEPTIDASES

THREONINE PEPTIDASES

GLUTAMIC PEPTIDASES

ASPARTIC PEPTIDASES

METALLO-PEPTIDASES

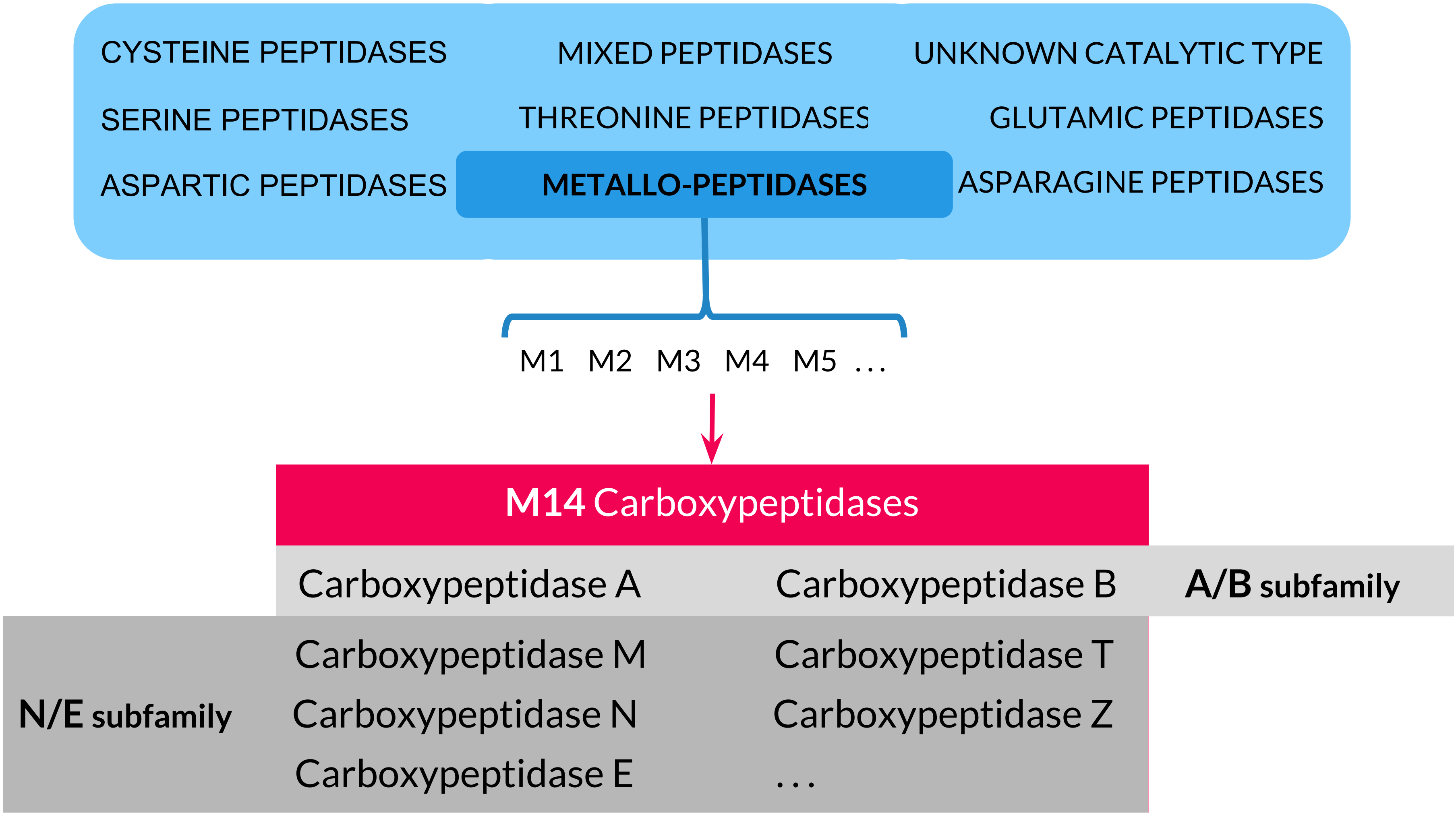
ASPARAGINE PEPTIDASES

M1 M2 M3 M4 M5 . . .

M14 Carboxypeptidases

Carboxypeptidase A	Carboxypeptidase B
Carboxypeptidase M	Carboxypeptidase T
Carboxypeptidase N	Carboxypeptidase Z
Carboxypeptidase E	. . .

Classification



CYSTEINE PEPTIDASES

SERINE PEPTIDASES

ASPARTIC PEPTIDASES

MIXED PEPTIDASES

THREONINE PEPTIDASES

METALLO-PEPTIDASES

UNKNOWN CATALYTIC TYPE

GLUTAMIC PEPTIDASES

ASPARAGINE PEPTIDASES

M1 M2 M3 M4 M5 . . .

M14 Carboxypeptidases

Carboxypeptidase A

Carboxypeptidase B

A/B subfamily

N/E subfamily

Carboxypeptidase M

Carboxypeptidase N

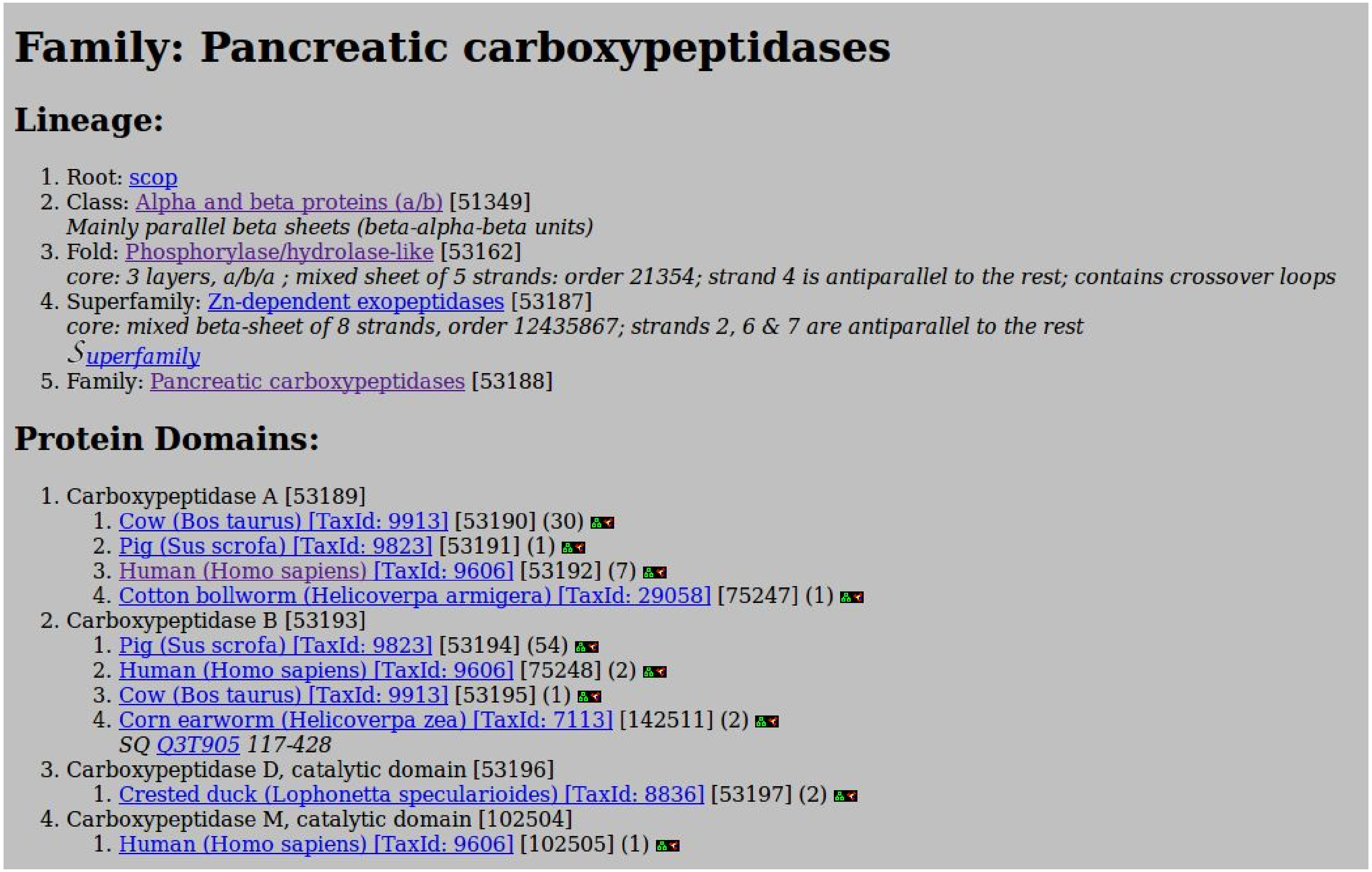
Carboxypeptidase E

Carboxypeptidase T

Carboxypeptidase Z

. . .

Classification SCOP classification



Family: Pancreatic carboxypeptidases

Lineage:

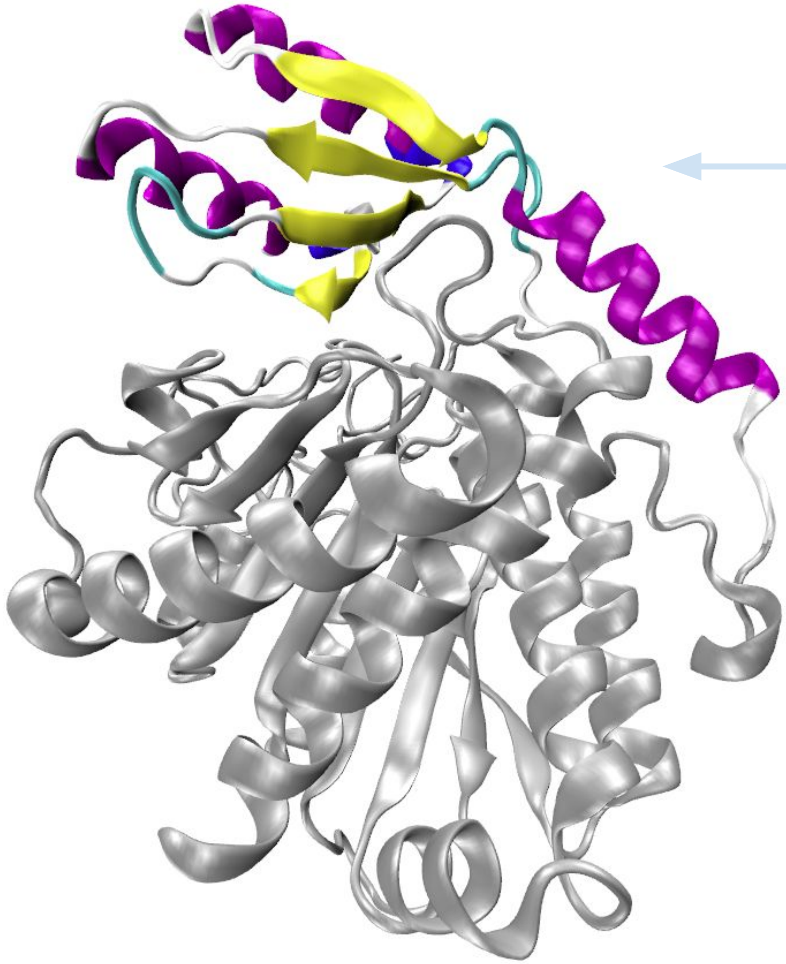
1. Root: scop
2. Class: Alpha and beta proteins (a/b) [51349]
   Mainly parallel beta sheets (beta-alpha-beta units)
3. Fold: Phosphorylase/hydrolase-like [53162]
   core: 3 layers, a/b/a ; mixed sheet of 5 strands: order 21354; strand 4 is antiparallel to the rest; contains crossover loops
4. Superfamily: Zn-dependent exopeptidases [53187]
   core: mixed beta-sheet of 8 strands, order 12435867; strands 2, 6 & 7 are antiparallel to the rest
   Superfamily
5. Family: Pancreatic carboxypeptidases [53188]

Protein Domains:

1. Carboxypeptidase A [53189]
   1. Cow (Bos taurus) [TaxId: 9913] [53190] (30)
   2. Pig (Sus scrofa) [TaxId: 9823] [53191] (1)
   3. Human (Homo sapiens) [TaxId: 9606] [53192] (7)
   4. Cotton bollworm (Helicoverpa armigera) [TaxId: 29058] [75247] (1)
2. Carboxypeptidase B [53193]
   1. Pig (Sus scrofa) [TaxId: 9823] [53194] (54)
   2. Human (Homo sapiens) [TaxId: 9606] [75248] (2)
   3. Cow (Bos taurus) [TaxId: 9913] [53195] (1)
   4. Corn earworm (Helicoverpa zea) [TaxId: 7113] [142511] (2)
      SQ Q3T905 117-428
3. Carboxypeptidase D, catalytic domain [53196]
   1. Crested duck (Lophonetta specularioides) [TaxId: 8836] [53197] (2)
4. Carboxypeptidase M, catalytic domain [102504]
   1. Human (Homo sapiens) [TaxId: 9606] [102505] (1)

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

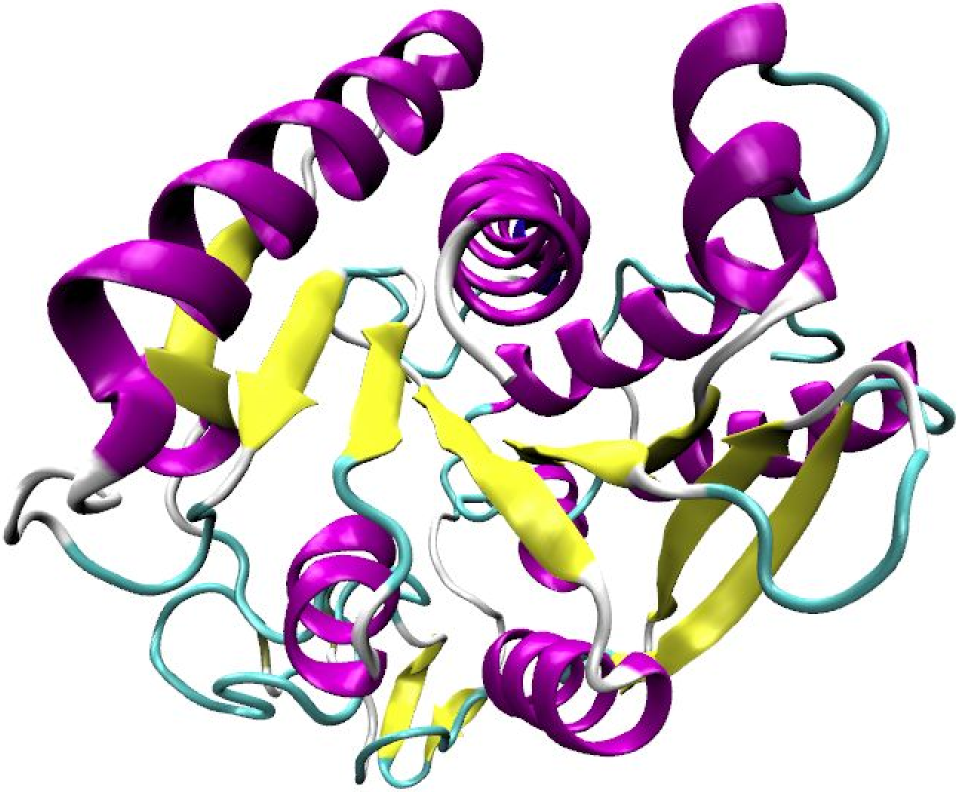
Procarboxypeptidase



propeptide

enzyme moiety

Rossmann fold-like



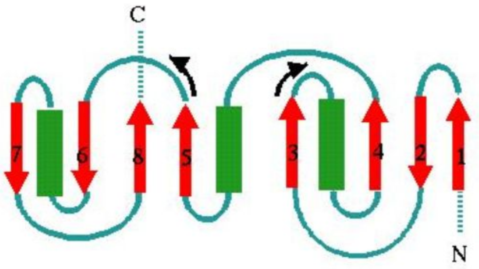
α / β

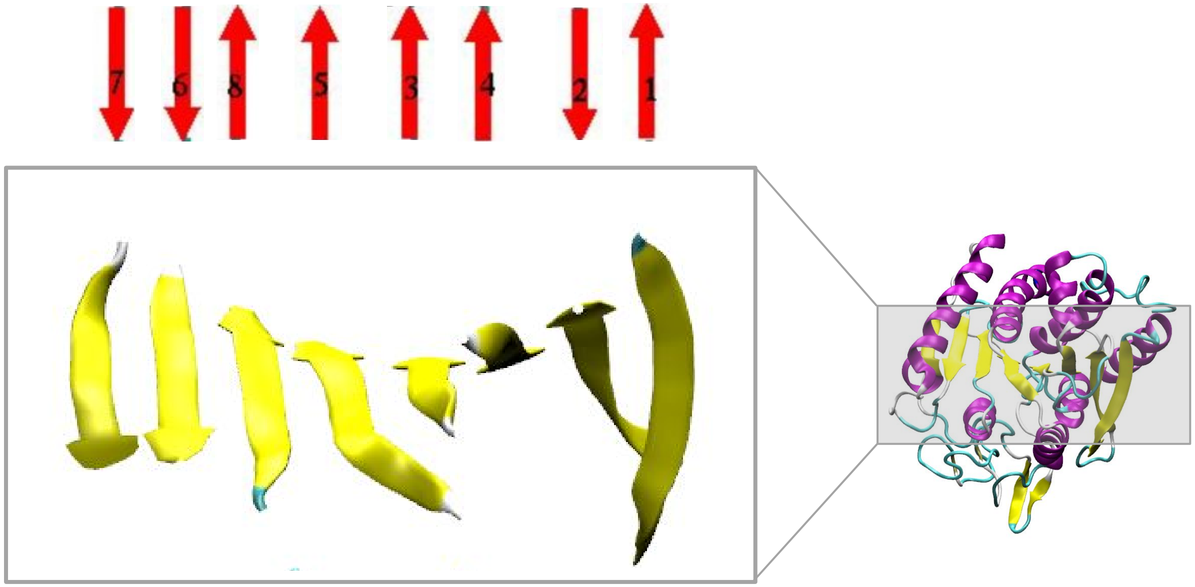
alpha helixes

beta sheets

alpha helixes

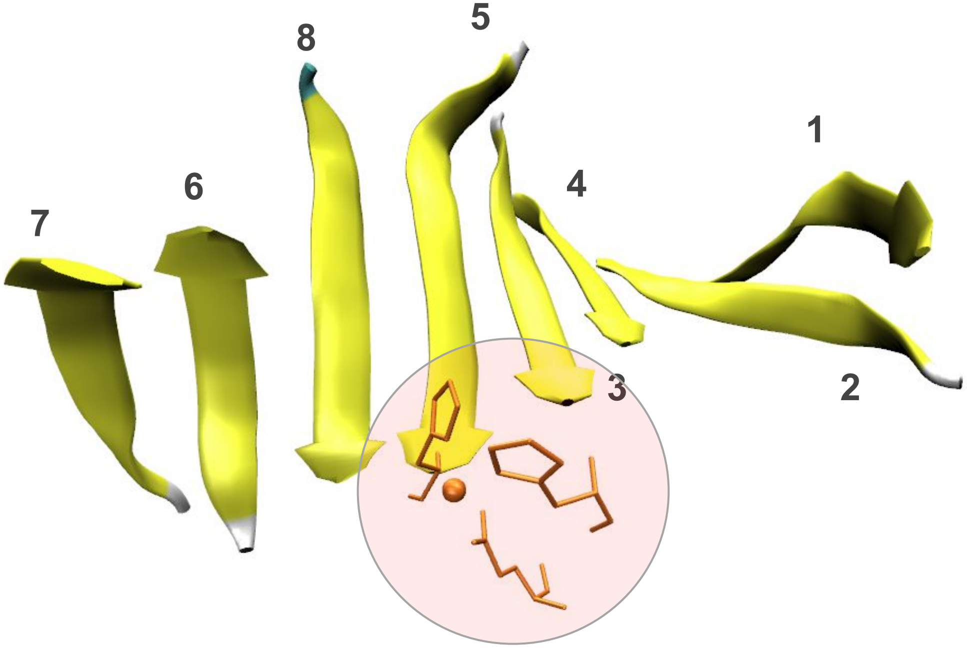
Rossmann fold-like

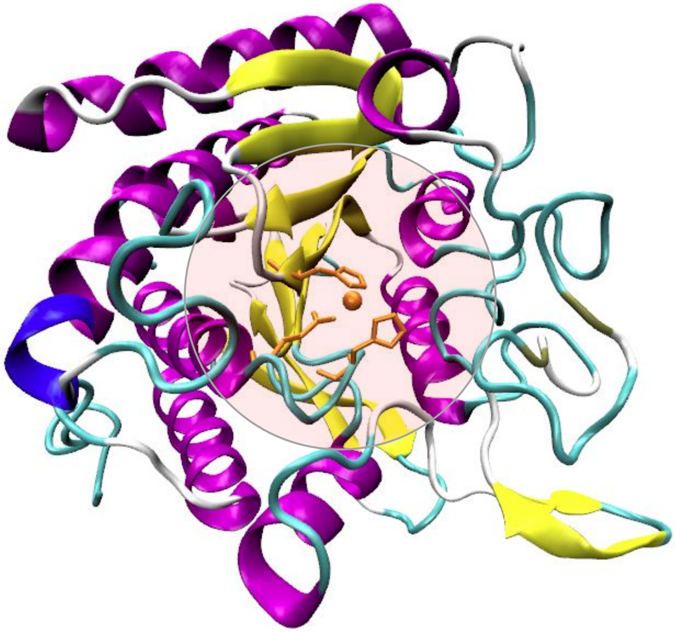




β-sheet

Active site

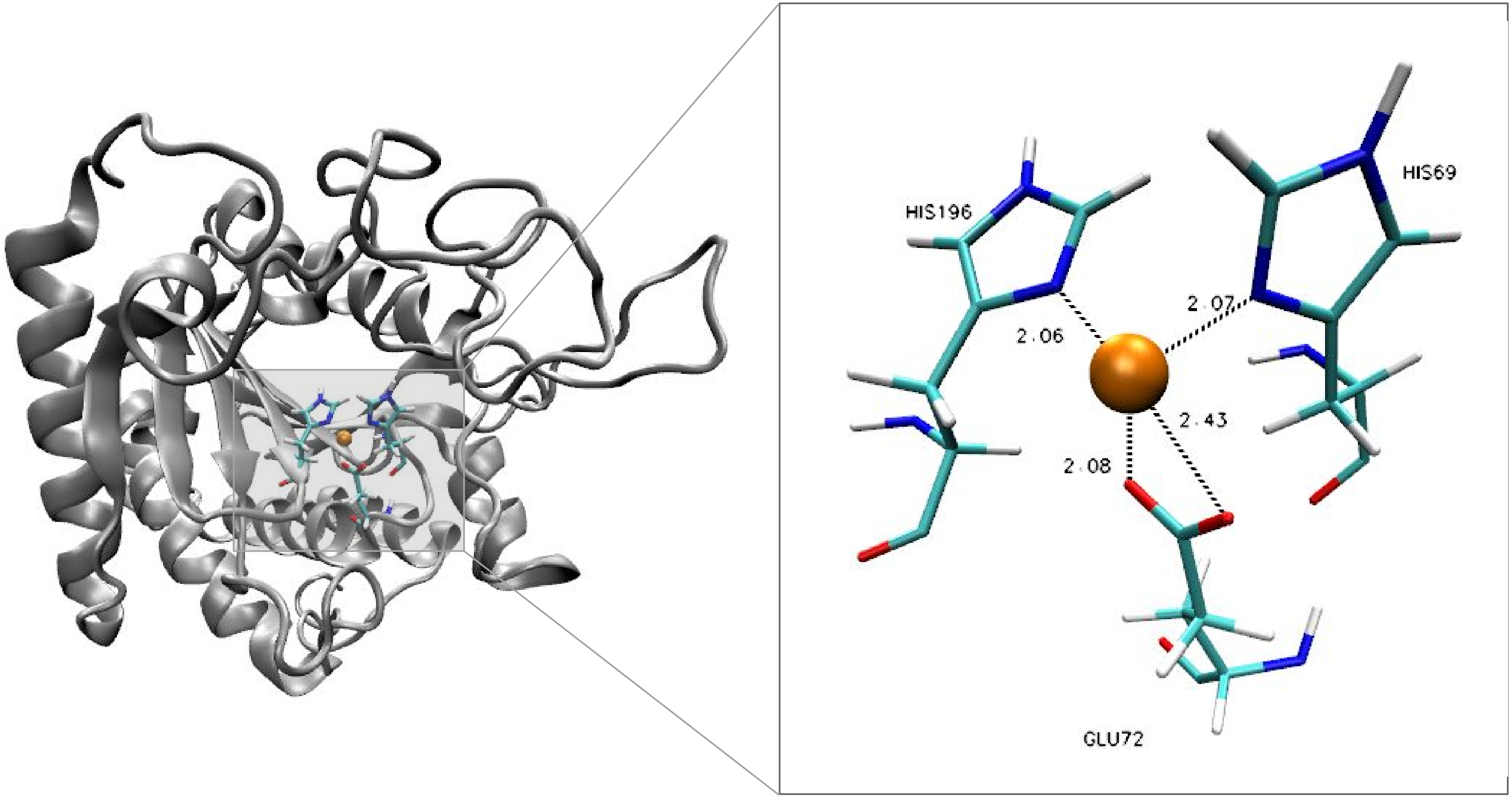




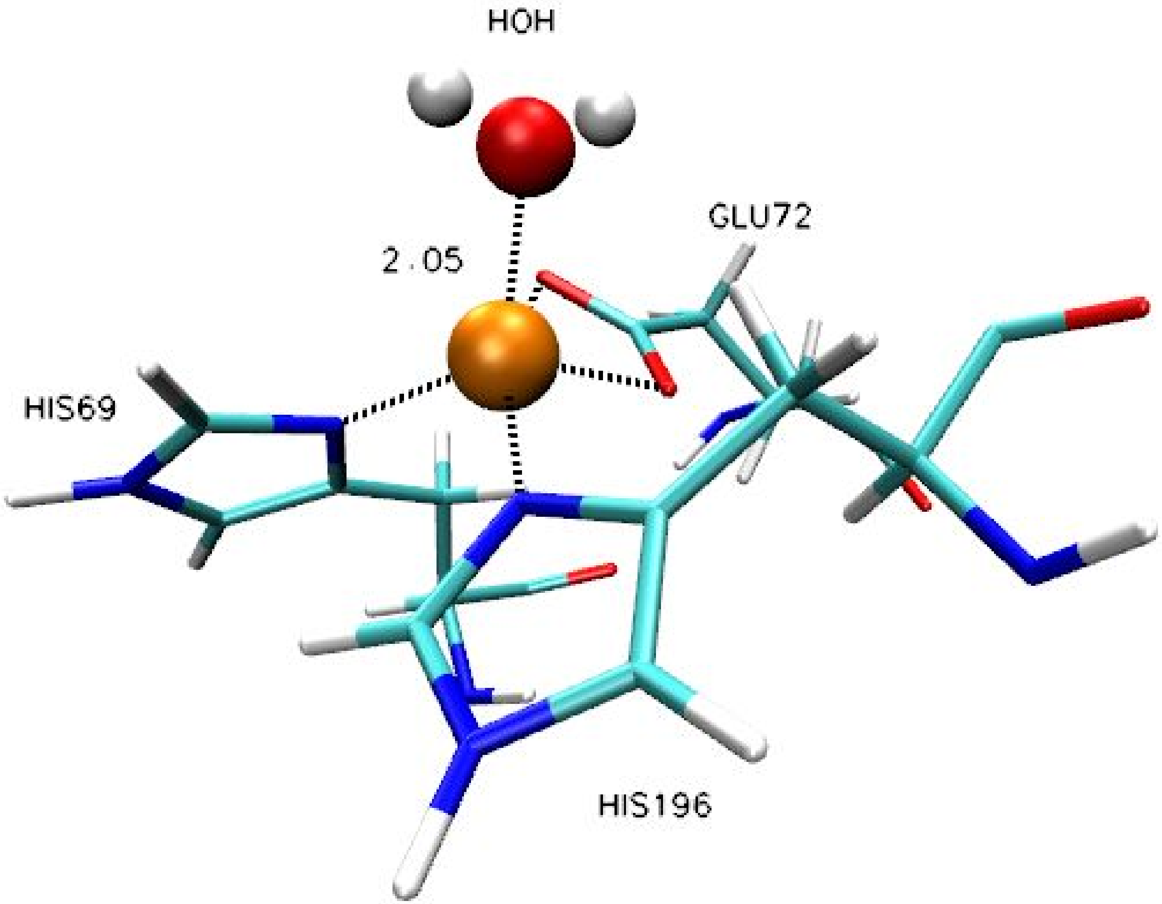
Active site

```
CBPA1_HUMAN_Homo_sapiens
CBPA1_MOUSE_Mus_musculus
CBPA1_BOVIN_Bos_taurus
B6A8G2_HELAM_Helicoverpa_armigera
CBPA1_RAT_Rattus_norvegicus
L5LE21_MYODS_Myotis_davidii
M7AG89_CHEMY_Chelonia_mydas
W6TYR4_ECHGR_Echinococcus_granulosus
A0A091GV45_9AVES_Cuculus_canorus
A0A093PIN1_PYGAD_Pygoscelis_adeliae
A0A091G9G5_9AVES_Cuculus_canorus
A0A1A7ZIR4_NOTFU_Nothobranchius_furzeri
A0A1A8ETY9_9TELE_=Nothobranchius_korthausae
C3KGU4_ANOFI_Anoplopoma_fimbria
L8WMV6_THACA_Thanatephorus_cucumeris
A0A0B0PAQ9_GOSAR_Gossypium_arboreum
A0A0V1DG89_TRIBR_Trichinella_britovi
K1QTU4_CRAGI_Crassostrea_gigas
L8JCQ3_9GAMM_Photobacterium_marinum
B0WJ40_CULQU_Culex_quinquefasciatus
A0A0K8UDD5_BACLA_Bactrocera_latifrons
A0A0N0PC47_PAPMA_Papilio_machaon
A0A0H5CKL5_9PSEU_Alloactinosynnema_sp.
A0A167BX44_9HYPO_Cordyceps_brongniartii
G3J5Q9_CORMM_Cordyceps_militaris
CBPA1_PIG_Sus_scrofa
```

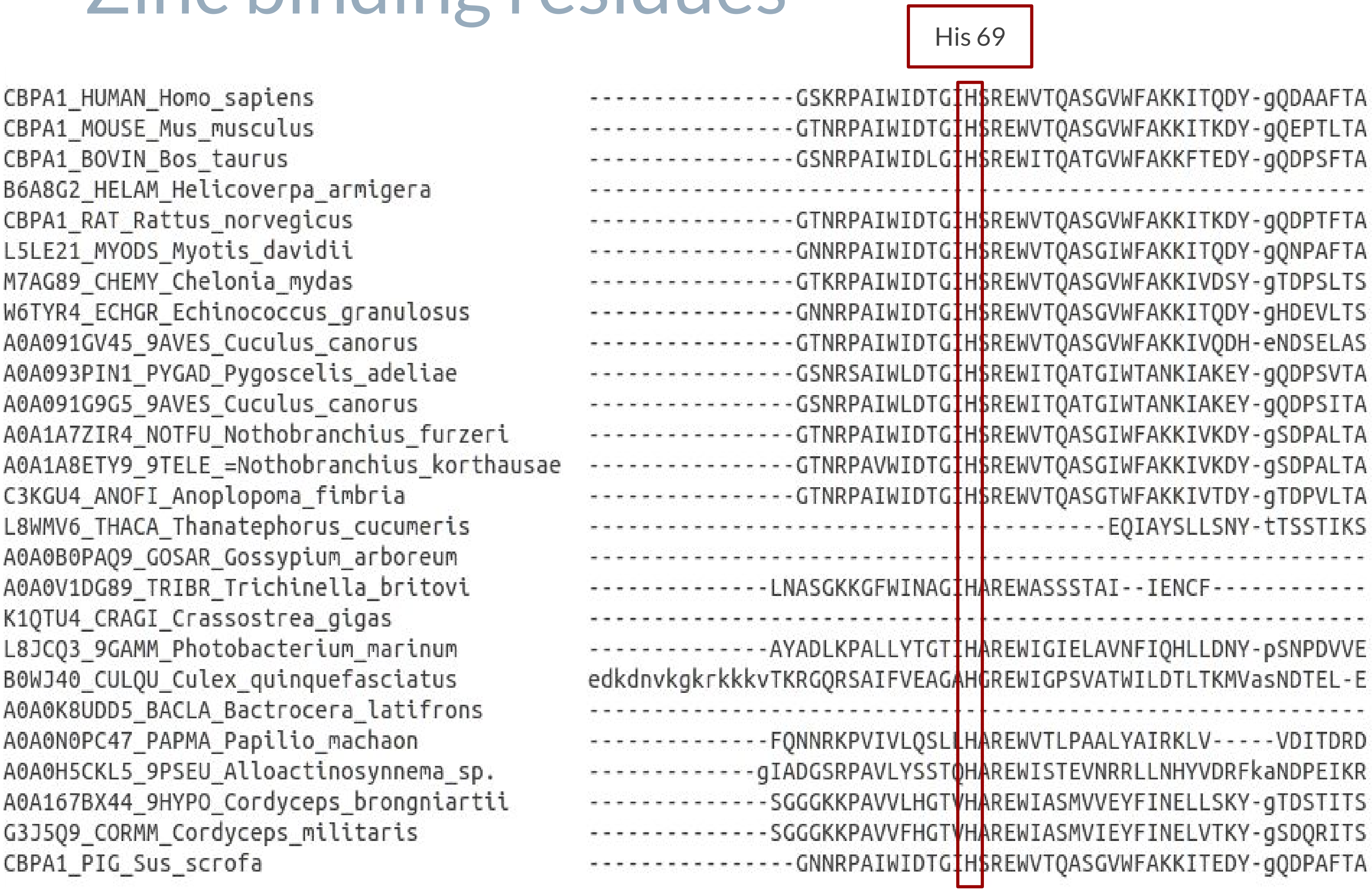
Zinc binding residues



Zinc binding residues Coordination sphere



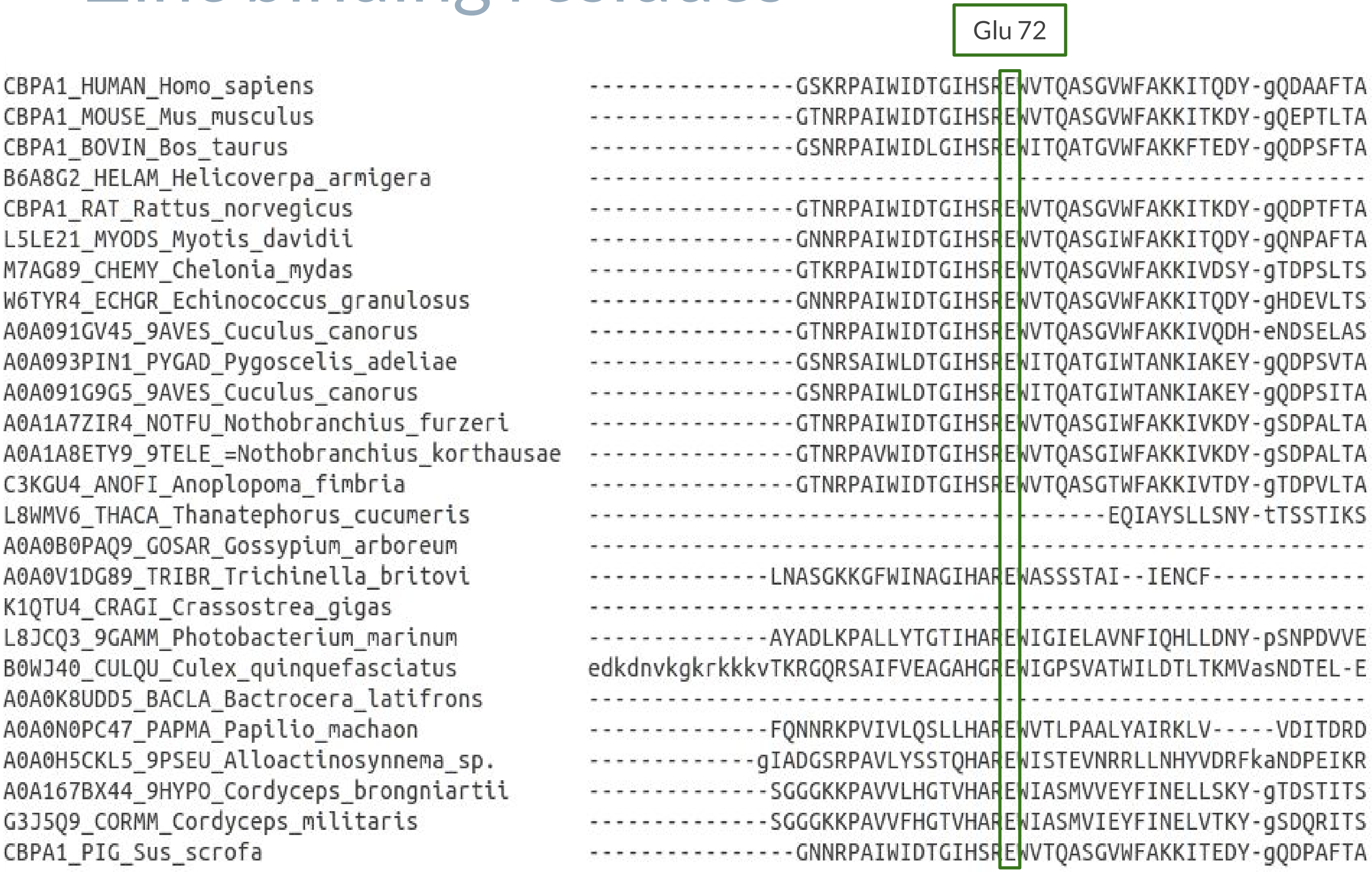
Zinc binding residues



His 69

```
CBPA1_HUMAN_Homo_sapiens                     ----------------GSKRPAIWIDTGIHSREWVTQASGVWFAKKITQDY-gQDAAFTA
CBPA1_MOUSE_Mus_musculus                     ----------------GTNRPAIWIDTGIHSREWVTQASGVWFAKKITKDY-gQEPTLTA
CBPA1_BOVIN_Bos_taurus                       ----------------GSNRPAIWIDLGIHSREWITQATGVWFAKKFTEDY-gQDPSFTA
B6A8G2_HELAM_Helicoverpa_armigera            ------------------------------------------------------------
CBPA1_RAT_Rattus_norvegicus                  ----------------GTNRPAIWIDTGIHSREWVTQASGVWFAKKITKDY-gQDPTFTA
L5LE21_MYODS_Myotis_davidii                  ----------------GNNRPAIWIDTGIHSREWVTQASGIWFAKKITQDY-gQNPAFTA
M7AG89_CHEMY_Chelonia_mydas                  ----------------GTKRPAIWIDTGIHSREWVTQASGVWFAKKIVDSY-gTDPSLTS
W6TYR4_ECHGR_Echinococcus_granulosus         ----------------GNNRPAIWIDTGIHSREWVTQASGVWFAKKITQDY-gHDEVLTS
A0A091GV45_9AVES_Cuculus_canorus             ----------------GTNRPAIWIDTGIHSREWVTQASGVWFAKKIVQDH-eNDSELAS
A0A093PIN1_PYGAD_Pygoscelis_adeliae          ----------------GSNRSAIWLDTGIHSREWITQATGIWTANKIAKEY-gQDPSVTA
A0A091G9G5_9AVES_Cuculus_canorus             ----------------GSNRPAIWLDTGIHSREWITQATGIWTANKIAKEY-gQDPSITA
A0A1A7ZIR4_NOTFU_Nothobranchius_furzeri      ----------------GTNRPAIWIDTGIHSREWVTQASGIWFAKKIVKDY-gSDPALTA
A0A1A8ETY9_9TELE_=Nothobranchius_korthausae  ----------------GTNRPAVWIDTGIHSREWVTQASGIWFAKKIVKDY-gSDPALTA
C3KGU4_ANOFI_Anoplopoma_fimbria              ----------------GTNRPAIWIDTGIHSREWVTQASGTWFAKKIVTDY-gTDPVLTA
L8WMV6_THACA_Thanatephorus_cucumeris         ----------------------------------------EQIAYSLLSNY-tTSSTIKS
A0A0B0PAQ9_GOSAR_Gossypium_arboreum          ------------------------------------------------------------
A0A0V1DG89_TRIBR_Trichinella_britovi         --------------LNASGKKGFWINAGIHAREWASSSTAI--IENCF------------
K1QTU4_CRAGI_Crassostrea_gigas               ------------------------------------------------------------
L8JCQ3_9GAMM_Photobacterium_marinum          --------------AYADLKPALLYTGTIHAREWIGIELAVNFIQHLLDNY-pSNPDVVE
B0WJ40_CULQU_Culex_quinquefasciatus          edkdnvkgkrkkkvTKRGQRSAIFVEAGAHGREWIGPSVATWILDTLTKMVasNDTEL-E
A0A0K8UDD5_BACLA_Bactrocera_latifrons        ------------------------------------------------------------
A0A0N0PC47_PAPMA_Papilio_machaon             --------------FQNNRKPVIVLQSLLHAREWVTLPAALYAIRKLV-----VDITDRD
A0A0H5CKL5_9PSEU_Alloactinosynnema_sp.       -------------gIADGSRPAVLYSSTQHAREWISTEVNRRLLNHYVDRFkaNDPEIKR
A0A167BX44_9HYPO_Cordyceps_brongniartii      --------------SGGGKKPAVVLHGTVHAREWIASMVVEYFINELLSKY-gTDSTITS
G3J5Q9_CORMM_Cordyceps_militaris             --------------SGGGKKPAVVFHGTVHAREWIASMVIEYFINELVTKY-gSDQRITS
CBPA1_PIG_Sus_scrofa                         ----------------GNNRPAIWIDTGIHSREWVTQASGVWFAKKITEDY-gQDPAFTA
```

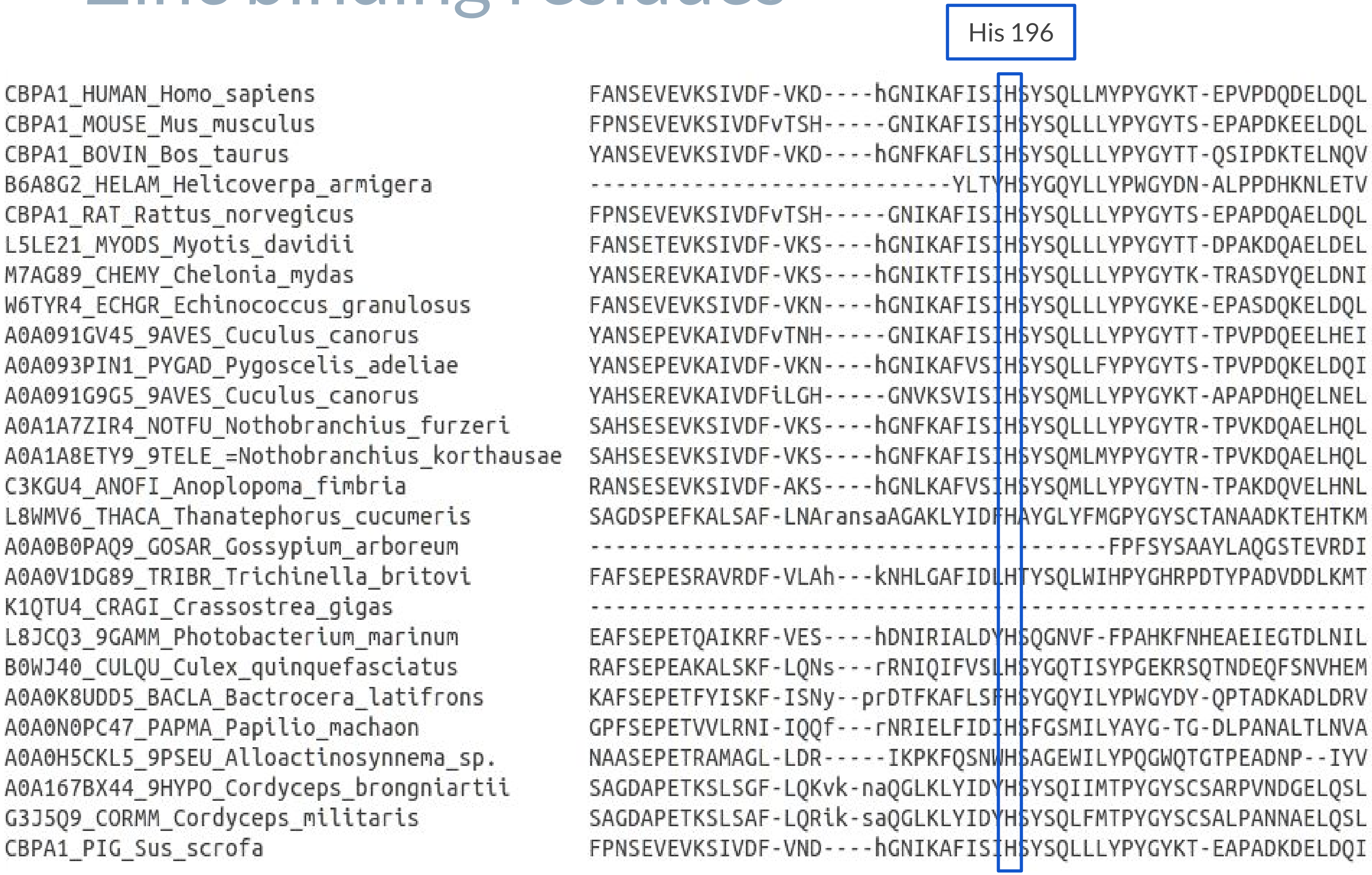
Zinc binding residues



Glu 72

```
CBPA1_HUMAN_Homo_sapiens                    ----------------GSKRPAIWIDTGIHSREWVTQASGVWFAKKITQDY-gQDAAFTA
CBPA1_MOUSE_Mus_musculus                    ----------------GTNRPAIWIDTGIHSREWVTQASGVWFAKKITKDY-gQEPTLTA
CBPA1_BOVIN_Bos_taurus                      ----------------GSNRPAIWIDLGIHSREWITQATGVWFAKKFTEDY-gQDPSFTA
B6A8G2_HELAM_Helicoverpa_armigera           ------------------------------------------------------------
CBPA1_RAT_Rattus_norvegicus                 ----------------GTNRPAIWIDTGIHSREWVTQASGVWFAKKITKDY-gQDPTFTA
L5LE21_MYODS_Myotis_davidii                 ----------------GNNRPAIWIDTGIHSREWVTQASGIWFAKKITQDY-gQNPAFTA
M7AG89_CHEMY_Chelonia_mydas                 ----------------GTKRPAIWIDTGIHSREWVTQASGVWFAKKIVDSY-gTDPSLTS
W6TYR4_ECHGR_Echinococcus_granulosus        ----------------GNNRPAIWIDTGIHSREWVTQASGVWFAKKITQDY-gHDEVLTS
A0A091GV45_9AVES_Cuculus_canorus            ----------------GTNRPAIWIDTGIHSREWVTQASGVWFAKKIVQDH-eNDSELAS
A0A093PIN1_PYGAD_Pygoscelis_adeliae         ----------------GSNRSAIWLDTGIHSREWITQATGIWTANKIAKEY-gQDPSVTA
A0A091G9G5_9AVES_Cuculus_canorus            ----------------GSNRPAIWLDTGIHSREWITQATGIWTANKIAKEY-gQDPSITA
A0A1A7ZIR4_NOTFU_Nothobranchius_furzeri     ----------------GTNRPAIWIDTGIHSREWVTQASGIWFAKKIVKDY-gSDPALTA
A0A1A8ETY9_9TELE_=Nothobranchius_korthausae ----------------GTNRPAVWIDTGIHSREWVTQASGIWFAKKIVKDY-gSDPALTA
C3KGU4_ANOFI_Anoplopoma_fimbria             ----------------GTNRPAIWIDTGIHSREWVTQASGTWFAKKIVTDY-gTDPVLTA
L8WMV6_THACA_Thanatephorus_cucumeris        ----------------------------------------EQIAYSLLSNY-tTSSTIKS
A0A0B0PAQ9_GOSAR_Gossypium_arboreum         ------------------------------------------------------------
A0A0V1DG89_TRIBR_Trichinella_britovi        --------------LNASGKKGFWINAGIHAREWASSSTAI--IENCF------------
K1QTU4_CRAGI_Crassostrea_gigas              ------------------------------------------------------------
L8JCQ3_9GAMM_Photobacterium_marinum         --------------AYADLKPALLYTGTIHAREWIGIELAVNFIQHLLDNY-pSNPDVVE
B0WJ40_CULQU_Culex_quinquefasciatus         edkdnvkgkrkkkvTKRGQRSAIFVEAGAHGREWIGPSVATWILDTLTKMVasNDTEL-E
A0A0K8UDD5_BACLA_Bactrocera_latifrons       ------------------------------------------------------------
A0A0N0PC47_PAPMA_Papilio_machaon            --------------FQNNRKPVIVLQSLLHAREWVTLPAALYAIRKLV-----VDITDRD
A0A0H5CKL5_9PSEU_Alloactinosynnema_sp.      -------------gIADGSRPAVLYSSTQHAREWISTEVNRRLLNHYVDRFkaNDPEIKR
A0A167BX44_9HYPO_Cordyceps_brongniartii     --------------SGGGKKPAVVLHGTVHAREWIASMVVEYFINELLSKY-gTDSTITS
G3J5Q9_CORMM_Cordyceps_militaris            --------------SGGGKKPAVVFHGTVHAREWIASMVIEYFINELVTKY-gSDQRITS
CBPA1_PIG_Sus_scrofa                        ----------------GNNRPAIWIDTGIHSREWVTQASGVWFAKKITEDY-gQDPAFTA
```

Zinc binding residues

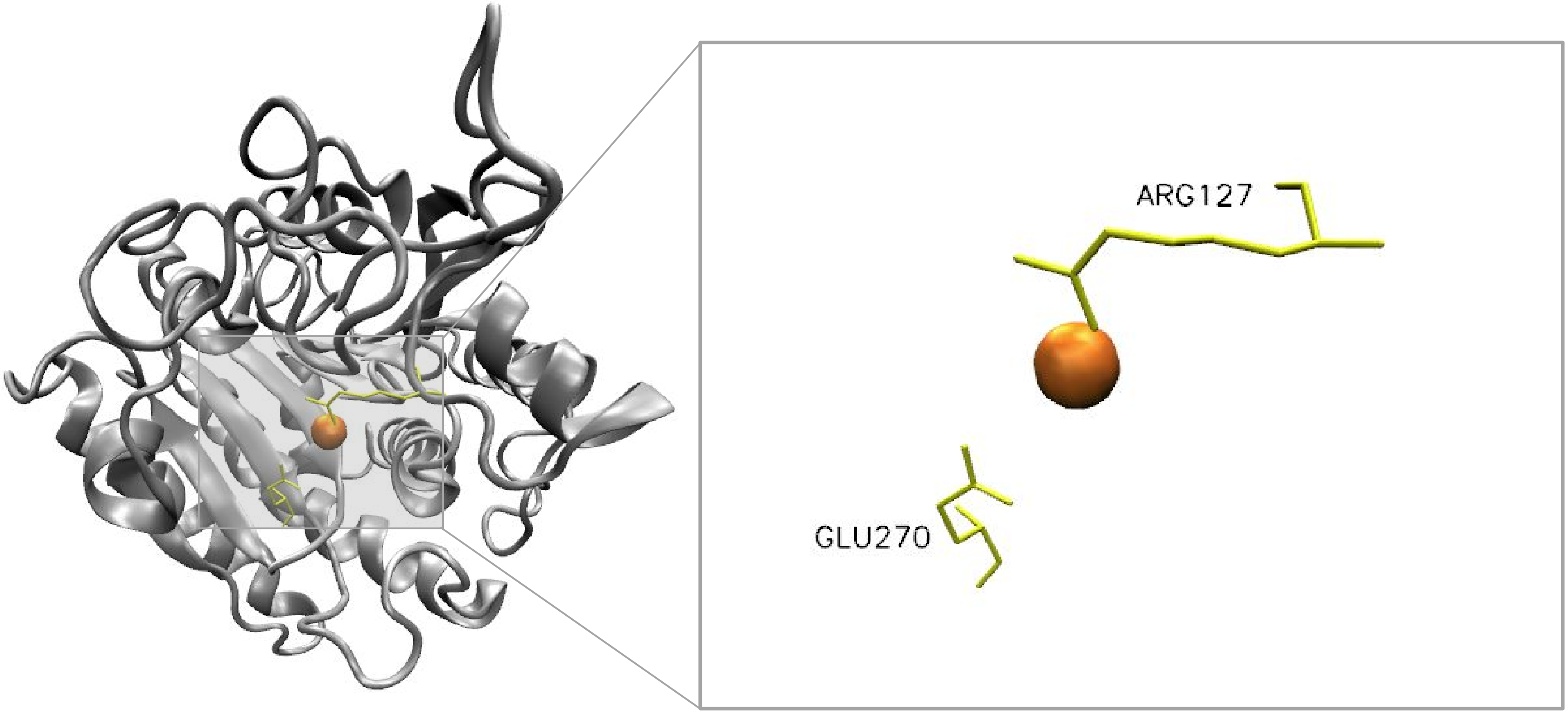


His 196

```
CBPA1_HUMAN_Homo_sapiens                     FANSEVEVKSIVDF-VKD----hGNIKAFISIHSYSQLLMYPYGYKT-EPVPDQDELDQL
CBPA1_MOUSE_Mus_musculus                     FPNSEVEVKSIVDFvTSH-----GNIKAFISIHSYSQLLLYPYGYTS-EPAPDKEELDQL
CBPA1_BOVIN_Bos_taurus                       YANSEVEVKSIVDF-VKD----hGNFKAFLSIHSYSQLLLYPYGYTT-QSIPDKTELNQV
B6A8G2_HELAM_Helicoverpa_armigera            ----------------------------YLTYHSYGQYLLYPWGYDN-ALPPDHKNLETV
CBPA1_RAT_Rattus_norvegicus                  FPNSEVEVKSIVDFvTSH-----GNIKAFISIHSYSQLLLYPYGYTS-EPAPDQAELDQL
L5LE21_MYODS_Myotis_davidii                  FANSETEVKSIVDF-VKS----hGNIKAFISIHSYSQLLLYPYGYTT-DPAKDQAELDEL
M7AG89_CHEMY_Chelonia_mydas                  YANSEREVKAIVDF-VKS----hGNIKTFISIHSYSQLLLYPYGYTK-TRASDYQELDNI
W6TYR4_ECHGR_Echinococcus_granulosus         FANSEVEVKSIVDF-VKN----hGNIKAFISIHSYSQLLLYPYGYKE-EPASDQKELDQL
A0A091GV45_9AVES_Cuculus_canorus             YANSEPEVKAIVDFvTNH-----GNIKAFISIHSYSQLLLYPYGYTT-TPVPDQEELHEI
A0A093PIN1_PYGAD_Pygoscelis_adeliae          YANSEPEVKAIVDF-VKN----hGNIKAFVSIHSYSQLLFYPYGYTS-TPVPDQKELDQI
A0A091G9G5_9AVES_Cuculus_canorus             YAHSEREVKAIVDFiLGH-----GNVKSVISIHSYSQMLLYPYGYKT-APAPDHQELNEL
A0A1A7ZIR4_NOTFU_Nothobranchius_furzeri      SAHSESEVKSIVDF-VKS----hGNFKAFISIHSYSQLLLYPYGYTR-TPVKDQAELHQL
A0A1A8ETY9_9TELE_=Nothobranchius_korthausae  SAHSESEVKSIVDF-VKS----hGNFKAFISIHSYSQMLMYPYGYTR-TPVKDQAELHQL
C3KGU4_ANOFI_Anoplopoma_fimbria              RANSESEVKSIVDF-AKS----hGNLKAFVSIHSYSQMLLYPYGYTN-TPAKDQVELHNL
L8WMV6_THACA_Thanatephorus_cucumeris         SAGDSPEFKALSAF-LNAransaAGAKLYIDFHAYGLYFMGPYGYSCTANAADKTEHTKM
A0A0B0PAQ9_GOSAR_Gossypium_arboreum          ----------------------------------------FPFSYSAAYLAQGSTEVRDI
A0A0V1DG89_TRIBR_Trichinella_britovi         FAFSEPESRAVRDF-VLAh---kNHLGAFIDLHTYSQLWIHPYGHRPDTYPADVDDLKMT
K1QTU4_CRAGI_Crassostrea_gigas               ------------------------------------------------------------
L8JCQ3_9GAMM_Photobacterium_marinum          EAFSEPETQAIKRF-VES----hDNIRIALDYHSQGNVF-FPAHKFNHEAEIEGTDLNIL
B0WJ40_CULQU_Culex_quinquefasciatus          RAFSEPEAKALSKF-LQNs---rRNIQIFVSLHSYGQTISYPGEKRSQTNDEQFSNVHEM
A0A0K8UDD5_BACLA_Bactrocera_latifrons        KAFSEPETFYISKF-ISNy--prDTFKAFLSFHSYGQYILYPWGYDY-QPTADKADLDRV
A0A0N0PC47_PAPMA_Papilio_machaon             GPFSEPETVVLRNI-IQQf---rNRIELFIDIHSFGSMILYAYG-TG-DLPANALTLNVA
A0A0H5CKL5_9PSEU_Alloactinosynnema_sp.       NAASEPETRAMAGL-LDR-----IKPKFQSNWHSAGEWILYPQGWQTGTPEADNP--IYV
A0A167BX44_9HYPO_Cordyceps_brongniartii      SAGDAPETKSLSGF-LQKvk-naQGLKLYIDYHSYSQIIMTPYGYSCSARPVNDGELQSL
G3J5Q9_CORMM_Cordyceps_militaris             SAGDAPETKSLSAF-LQRik-saQGLKLYIDYHSYSQLFMTPYGYSCSALPANNAELQSL
CBPA1_PIG_Sus_scrofa                         FPNSEVEVKSIVDF-VND----hGNIKAFISIHSYSQLLLYPYGYKT-EAPADKDELDQI
```

Active site

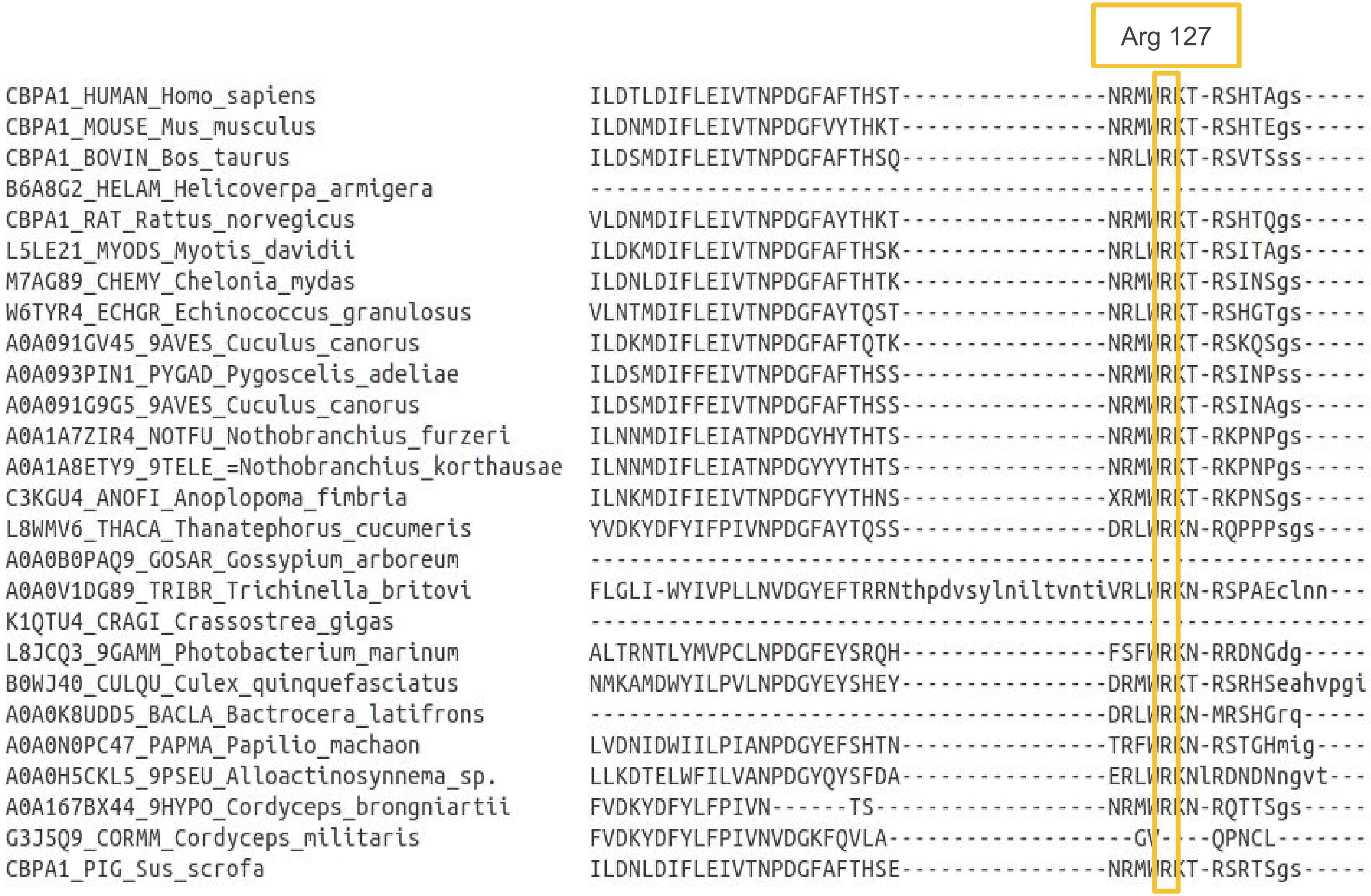
Subsite 1



S1	**Arg 127** **Glu 270**

Active site

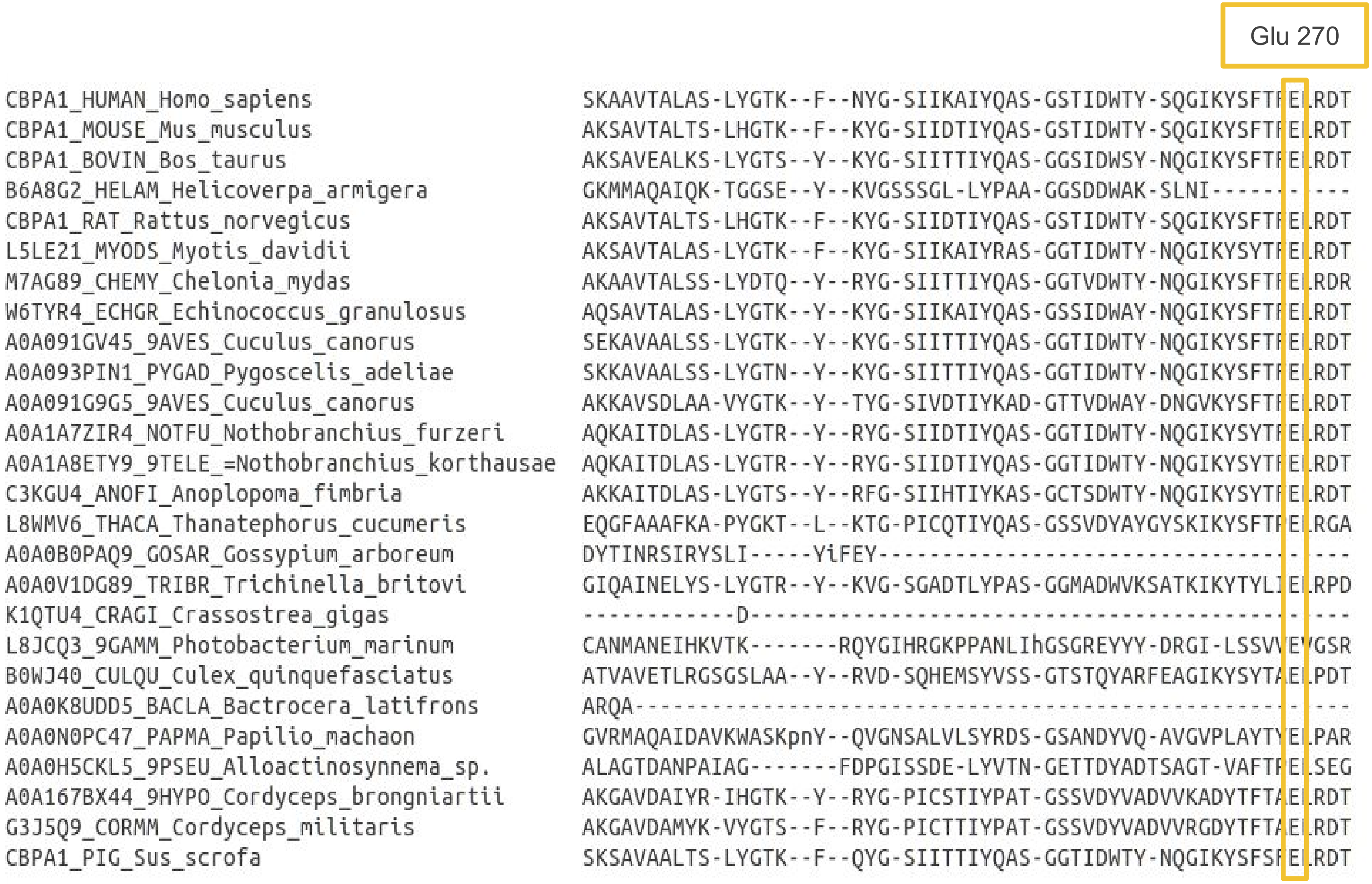
Subsite 1


Arg 127

```
CBPA1_HUMAN_Homo_sapiens                     ILDTLDIFLEIVTNPDGFAFTHST----------------NRMWRKT-RSHTAgs-----
CBPA1_MOUSE_Mus_musculus                     ILDNMDIFLEIVTNPDGFVYTHKT----------------NRMWRKT-RSHTEgs-----
CBPA1_BOVIN_Bos_taurus                       ILDSMDIFLEIVTNPDGFAFTHSQ----------------NRLWRKT-RSVTSss-----
B6A8G2_HELAM_Helicoverpa_armigera            ------------------------------------------------------------
CBPA1_RAT_Rattus_norvegicus                  VLDNMDIFLEIVTNPDGFAYTHKT----------------NRMWRKT-RSHTQgs-----
L5LE21_MYODS_Myotis_davidii                  ILDKMDIFLEIVTNPDGFAFTHSK----------------NRLWRKT-RSITAgs-----
M7AG89_CHEMY_Chelonia_mydas                  ILDNLDIFLEIVTNPDGFAFTHTK----------------NRMWRKT-RSINSgs-----
W6TYR4_ECHGR_Echinococcus_granulosus         VLNTMDIFLEIVTNPDGFAYTQST----------------NRLWRKT-RSHGTgs-----
A0A091GV45_9AVES_Cuculus_canorus             ILDKMDIFLEIVTNPDGFAFTQTK----------------NRMWRKT-RSKQSgs-----
A0A093PIN1_PYGAD_Pygoscelis_adeliae          ILDSMDIFFEIVTNPDGFAFTHSS----------------NRMWRKT-RSINPss-----
A0A091G9G5_9AVES_Cuculus_canorus             ILDSMDIFFEIVTNPDGFAFTHSS----------------NRMWRKT-RSINAgs-----
A0A1A7ZIR4_NOTFU_Nothobranchius_furzeri      ILNNMDIFLEIATNPDGYHYTHTS----------------NRMWRKT-RKPNPgs-----
A0A1A8ETY9_9TELE_=Nothobranchius_korthausae  ILNNMDIFLEIATNPDGYYYTHTS----------------NRMWRKT-RKPNPgs-----
C3KGU4_ANOFI_Anoplopoma_fimbria              ILNKMDIFIEIVTNPDGFYYTHNS----------------XRMWRKT-RKPNSgs-----
L8WMV6_THACA_Thanatephorus_cucumeris         YVDKYDFYIFPIVNPDGFAYTQSS----------------DRLWRKN-RQPPPsgs----
A0A0B0PAQ9_GOSAR_Gossypium_arboreum          ------------------------------------------------------------
A0A0V1DG89_TRIBR_Trichinella_britovi         FLGLI-WYIVPLLNVDGYEFTRRNthpdvsylniltvntiVRLWRKN-RSPAEclnn---
K1QTU4_CRAGI_Crassostrea_gigas               ------------------------------------------------------------
L8JCQ3_9GAMM_Photobacterium_marinum          ALTRNTLYMVPCLNPDGFEYSRQH----------------FSFWRKN-RRDNGdg-----
B0WJ40_CULQU_Culex_quinquefasciatus          NMKAMDWYILPVLNPDGYEYSHEY----------------DRMWRKT-RSRHSeahvpgi
A0A0K8UDD5_BACLA_Bactrocera_latifrons        ----------------------------------------DRLWRKN-MRSHGrq-----
A0A0N0PC47_PAPMA_Papilio_machaon             LVDNIDWIILPIANPDGYEFSHTN----------------TRFWRKN-RSTGHmig----
A0A0H5CKL5_9PSEU_Alloactinosynnema_sp.       LLKDTELWFILVANPDGYQYSFDA----------------ERLWRKNlRDNDNngvt---
A0A167BX44_9HYPO_Cordyceps_brongniartii      FVDKYDFYLFPIVN------TS------------------NRMWRKN-RQTTSgs-----
G3J5Q9_CORMM_Cordyceps_militaris             FVDKYDFYLFPIVNVDGKFQVLA-------------------GV----QPNCL-------
CBPA1_PIG_Sus_scrofa                         ILDNLDIFLEIVTNPDGFAFTHSE----------------NRMWRKT-RSRTSgs-----
```

Active site

Subsite 1

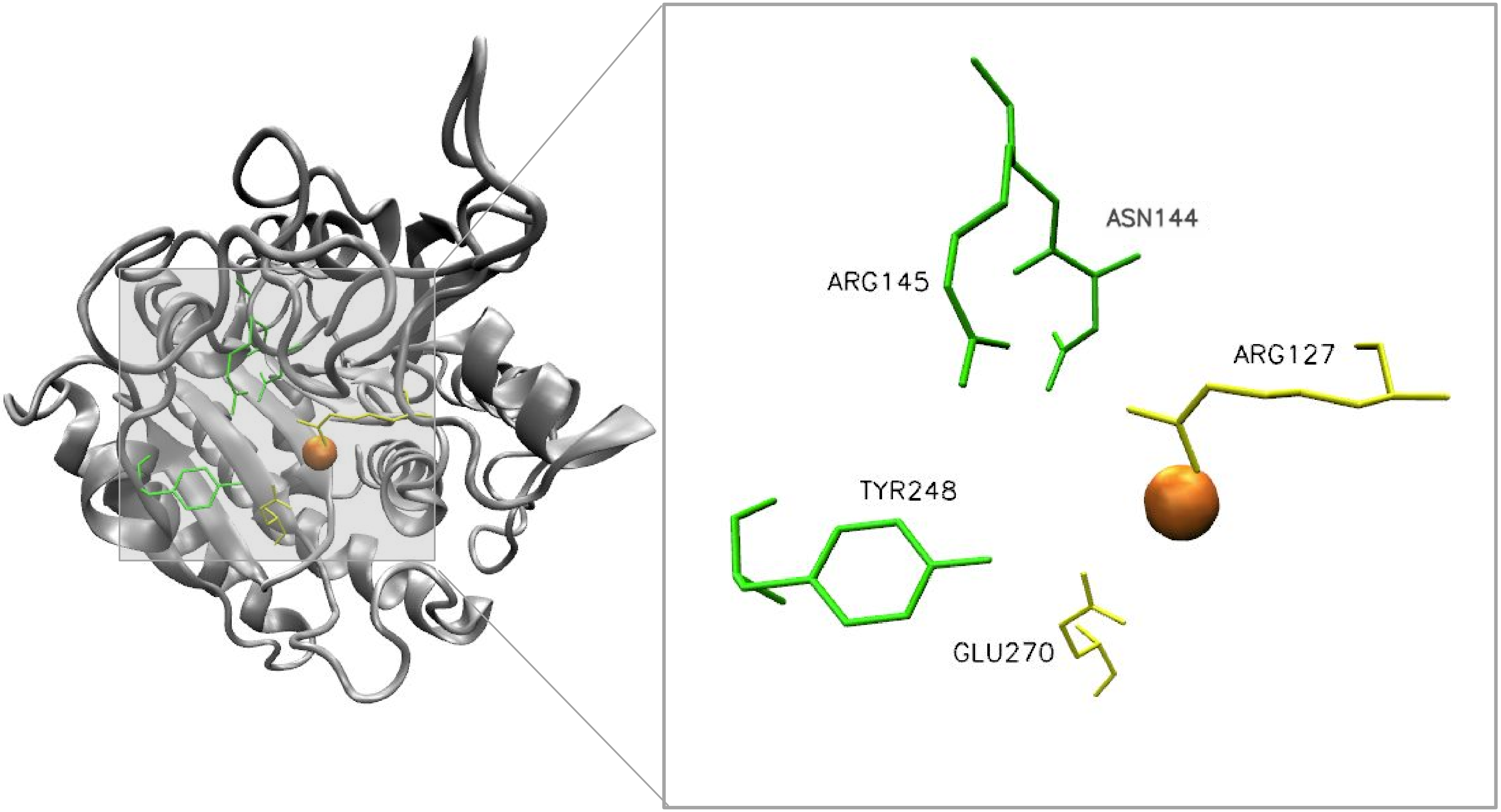


Glu 270

```
CBPA1_HUMAN_Homo_sapiens                    SKAAVTALAS-LYGTK--F--NYG-SIIKAIYQAS-GSTIDWTY-SQGIKYSFTFELRDT
CBPA1_MOUSE_Mus_musculus                    AKSAVTALTS-LHGTK--F--KYG-SIIDTIYQAS-GSTIDWTY-SQGIKYSFTFELRDT
CBPA1_BOVIN_Bos_taurus                      AKSAVEALKS-LYGTS--Y--KYG-SIITTIYQAS-GGSIDWSY-NQGIKYSFTFELRDT
B6A8G2_HELAM_Helicoverpa_armigera           GKMMAQAIQK-TGGSE--Y--KVGSSSGL-LYPAA-GGSDDWAK-SLNI-----------
CBPA1_RAT_Rattus_norvegicus                 AKSAVTALTS-LHGTK--F--KYG-SIIDTIYQAS-GSTIDWTY-SQGIKYSFTFELRDT
L5LE21_MYODS_Myotis_davidii                 AKSAVTALAS-LYGTK--F--KYG-SIIKAIYRAS-GGTIDWTY-NQGIKYSYTFELRDT
M7AG89_CHEMY_Chelonia_mydas                 AKAAVTALSS-LYDTQ--Y--RYG-SIITTIYQAS-GGTVDWTY-NQGIKYSFTFELRDR
W6TYR4_ECHGR_Echinococcus_granulosus        AQSAVTALAS-LYGTK--Y--KYG-SIIKAIYQAS-GSSIDWAY-NQGIKYSFTFELRDT
A0A091GV45_9AVES_Cuculus_canorus            SEKAVAALSS-LYGTK--Y--KYG-SIITTIYQAS-GGTIDWTY-NQGIKYSFTFELRDT
A0A093PIN1_PYGAD_Pygoscelis_adeliae         SKKAVAALSS-LYGTN--Y--KYG-SIITTIYQAS-GGTIDWTY-NQGIKYSFTFELRDT
A0A091G9G5_9AVES_Cuculus_canorus            AKKAVSDLAA-VYGTK--Y--TYG-SIVDTIYKAD-GTTVDWAY-DNGVKYSFTFELRDT
A0A1A7ZIR4_NOTFU_Nothobranchius_furzeri     AQKAITDLAS-LYGTR--Y--RYG-SIIDTIYQAS-GGTIDWTY-NQGIKYSYTFELRDT
A0A1A8ETY9_9TELE_=Nothobranchius_korthausae AQKAITDLAS-LYGTR--Y--RYG-SIIDTIYQAS-GGTIDWTY-NQGIKYSYTFELRDT
C3KGU4_ANOFI_Anoplopoma_fimbria             AKKAITDLAS-LYGTS--Y--RFG-SIIHTIYKAS-GCTSDWTY-NQGIKYSYTFELRDT
L8WMV6_THACA_Thanatephorus_cucumeris        EQGFAAAFKA-PYGKT--L--KTG-PICQTIYQAS-GSSVDYAYGYSKIKYSFTPELRGA
A0A0B0PAQ9_GOSAR_Gossypium_arboreum         DYTINRSIRYSLI-----YiFEY-------------------------------------
A0A0V1DG89_TRIBR_Trichinella_britovi        GIQAINELYS-LYGTR--Y--KVG-SGADTLYPAS-GGMADWVKSATKIKYTYLIELRPD
K1QTU4_CRAGI_Crassostrea_gigas              ------------D-----------------------------------------------
L8JCQ3_9GAMM_Photobacterium_marinum         CANMANEIHKVTK-------RQYGIHRGKPPANLIhGSGREYYY-DRGI-LSSVVEVGSR
B0WJ40_CULQU_Culex_quinquefasciatus         ATVAVETLRGSGSLAA--Y--RVD-SQHEMSYVSS-GTSTQYARFEAGIKYSYTAELPDT
A0A0K8UDD5_BACLA_Bactrocera_latifrons       ARQA--------------------------------------------------------
A0A0N0PC47_PAPMA_Papilio_machaon            GVRMAQAIDAVKWASKpnY--QVGNSALVLSYRDS-GSANDYVQ-AVGVPLAYTYELPAR
A0A0H5CKL5_9PSEU_Alloactinosynnema_sp.      ALAGTDANPAIAG-------FDPGISSDE-LYVTN-GETTDYADTSAGT-VAFTPELSEG
A0A167BX44_9HYPO_Cordyceps_brongniartii     AKGAVDAIYR-IHGTK--Y--RYG-PICSTIYPAT-GSSVDYVADVVKADYTFTAELRDT
G3J5Q9_CORMM_Cordyceps_militaris            AKGAVDAMYK-VYGTS--F--RYG-PICTTIYPAT-GSSVDYVADVVRGDYTFTAELRDT
CBPA1_PIG_Sus_scrofa                        SKSAVAALTS-LYGTK--F--QYG-SIITTIYQAS-GGTIDWTY-NQGIKYSFSFELRDT
```

Active site

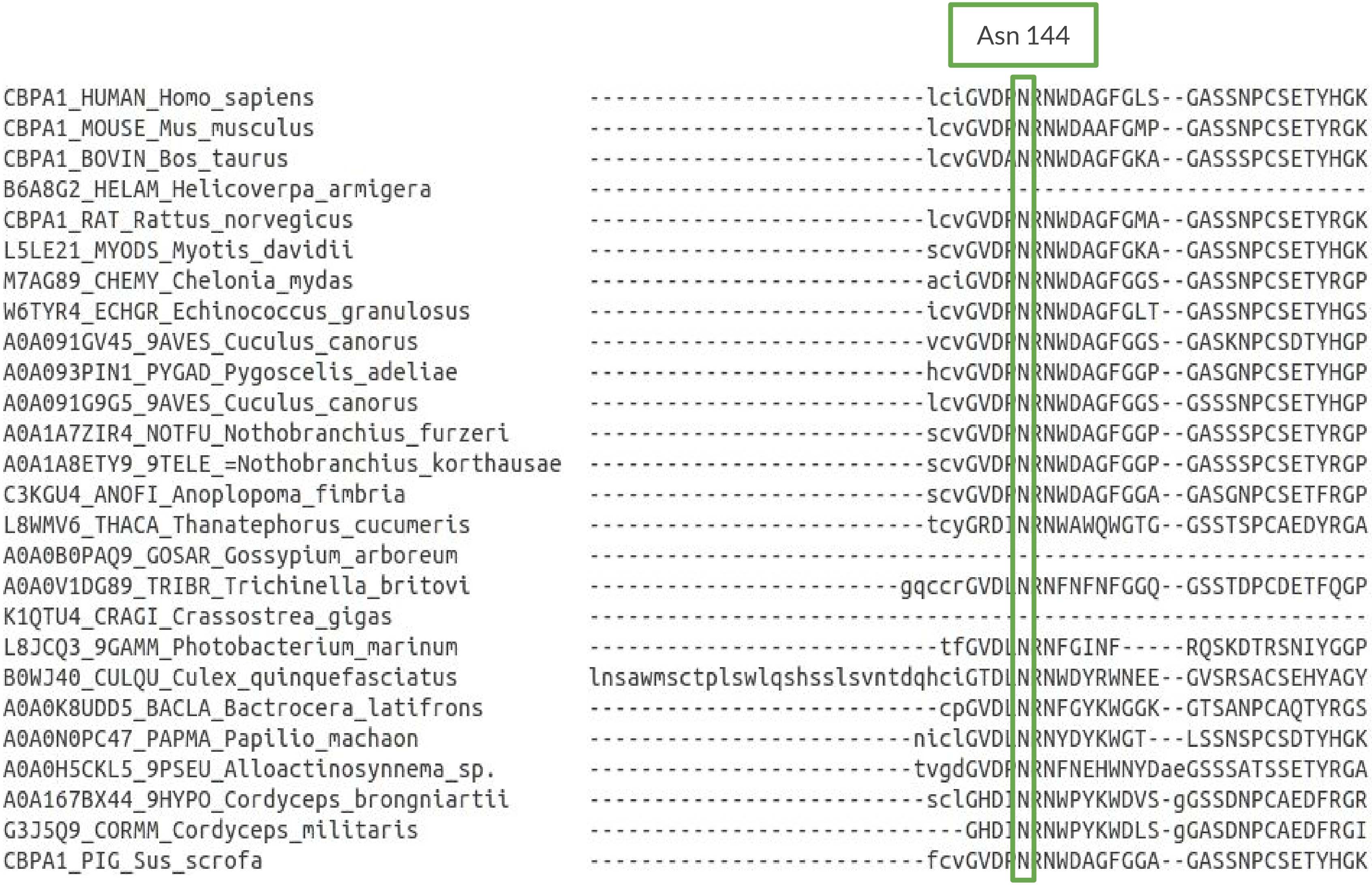
Subsite 1’



S1	Arg 127 Glu 270	S1’	Asn 144 Arg 145 Tyr 248

Active site

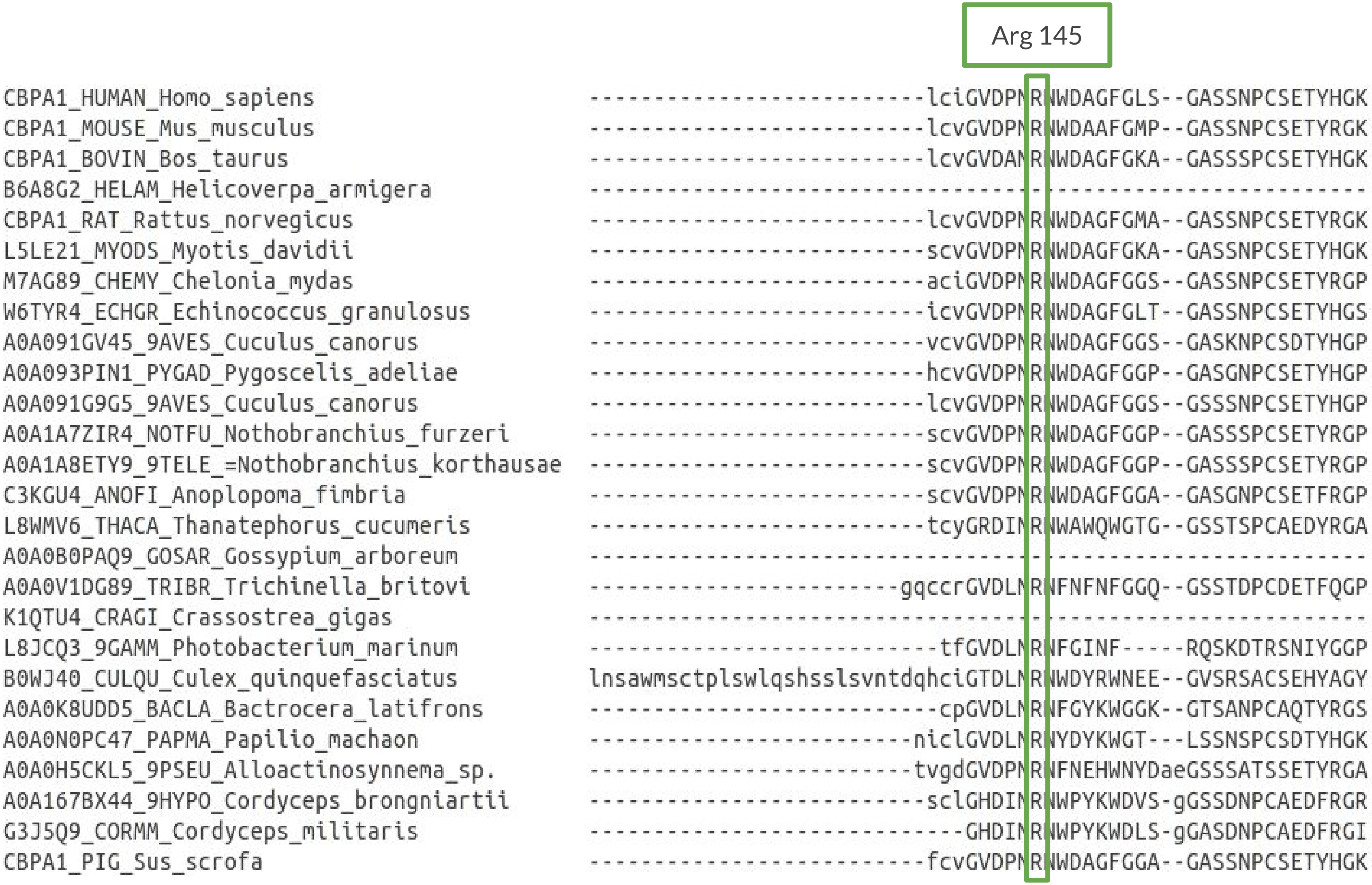
Subsite 1’



```
CBPA1_HUMAN_Homo_sapiens                     --------------------------lciGVDPNRNWDAGFGLS--GASSNPCSETYHGK
CBPA1_MOUSE_Mus_musculus                     --------------------------lcvGVDPNRNWDAAFGMP--GASSNPCSETYRGK
CBPA1_BOVIN_Bos_taurus                       --------------------------lcvGVDANRNWDAGFGKA--GASSSPCSETYHGK
B6A8G2_HELAM_Helicoverpa_armigera            ------------------------------------------------------------
CBPA1_RAT_Rattus_norvegicus                  --------------------------lcvGVDPNRNWDAGFGMA--GASSNPCSETYRGK
L5LE21_MYODS_Myotis_davidii                  --------------------------scvGVDPNRNWDAGFGKA--GASSNPCSETYHGK
M7AG89_CHEMY_Chelonia_mydas                  --------------------------aciGVDPNRNWDAGFGGS--GASSNPCSETYRGP
W6TYR4_ECHGR_Echinococcus_granulosus         --------------------------icvGVDPNRNWDAGFGLT--GASSNPCSETYHGS
A0A091GV45_9AVES_Cuculus_canorus             --------------------------vcvGVDPNRNWDAGFGGS--GASKNPCSDTYHGP
A0A093PIN1_PYGAD_Pygoscelis_adeliae          --------------------------hcvGVDPNRNWDAGFGGP--GASGNPCSETYHGP
A0A091G9G5_9AVES_Cuculus_canorus             --------------------------lcvGVDPNRNWDAGFGGS--GSSSNPCSETYHGP
A0A1A7ZIR4_NOTFU_Nothobranchius_furzeri      --------------------------scvGVDPNRNWDAGFGGP--GASSSPCSETYRGP
A0A1A8ETY9_9TELE_=Nothobranchius_korthausae  --------------------------scvGVDPNRNWDAGFGGP--GASSSPCSETYRGP
C3KGU4_ANOFI_Anoplopoma_fimbria              --------------------------scvGVDPNRNWDAGFGGA--GASGNPCSETFRGP
L8WMV6_THACA_Thanatephorus_cucumeris         --------------------------tcyGRDINRNWAWQWGTG--GSSTSPCAEDYRGA
A0A0B0PAQ9_GOSAR_Gossypium_arboreum          ------------------------------------------------------------
A0A0V1DG89_TRIBR_Trichinella_britovi         ------------------------gqccrGVDLNRNFNFNFGGQ--GSSTDPCDETFQGP
K1QTU4_CRAGI_Crassostrea_gigas               ------------------------------------------------------------
L8JCQ3_9GAMM_Photobacterium_marinum          ---------------------------tfGVDLNRNFGINF-----RQSKDTRSNIYGGP
B0WJ40_CULQU_Culex_quinquefasciatus          lnsawmsctplswlqshsslsvntdqhciGTDLNRNWDYRWNEE--GVSRSACSEHYAGY
A0A0K8UDD5_BACLA_Bactrocera_latifrons        ---------------------------cpGVDLNRNFGYKWGGK--GTSANPCAQTYRGS
A0A0N0PC47_PAPMA_Papilio_machaon             -------------------------niclGVDLNRNYDYKWGT---LSSNSPCSDTYHGK
A0A0H5CKL5_9PSEU_Alloactinosynnema_sp.       -------------------------tvgdGVDPNRNFNEHWNYDaeGSSSATSSETYRGA
A0A167BX44_9HYPO_Cordyceps_brongniartii      --------------------------sclGHDINRNWPYKWDVS-gGSSDNPCAEDFRGR
G3J5Q9_CORMM_Cordyceps_militaris             -----------------------------GHDINRNWPYKWDLS-gGASDNPCAEDFRGI
CBPA1_PIG_Sus_scrofa                         --------------------------fcvGVDPNRNWDAGFGGA--GASSNPCSETYHGK
```

Active site

Subsite 1’

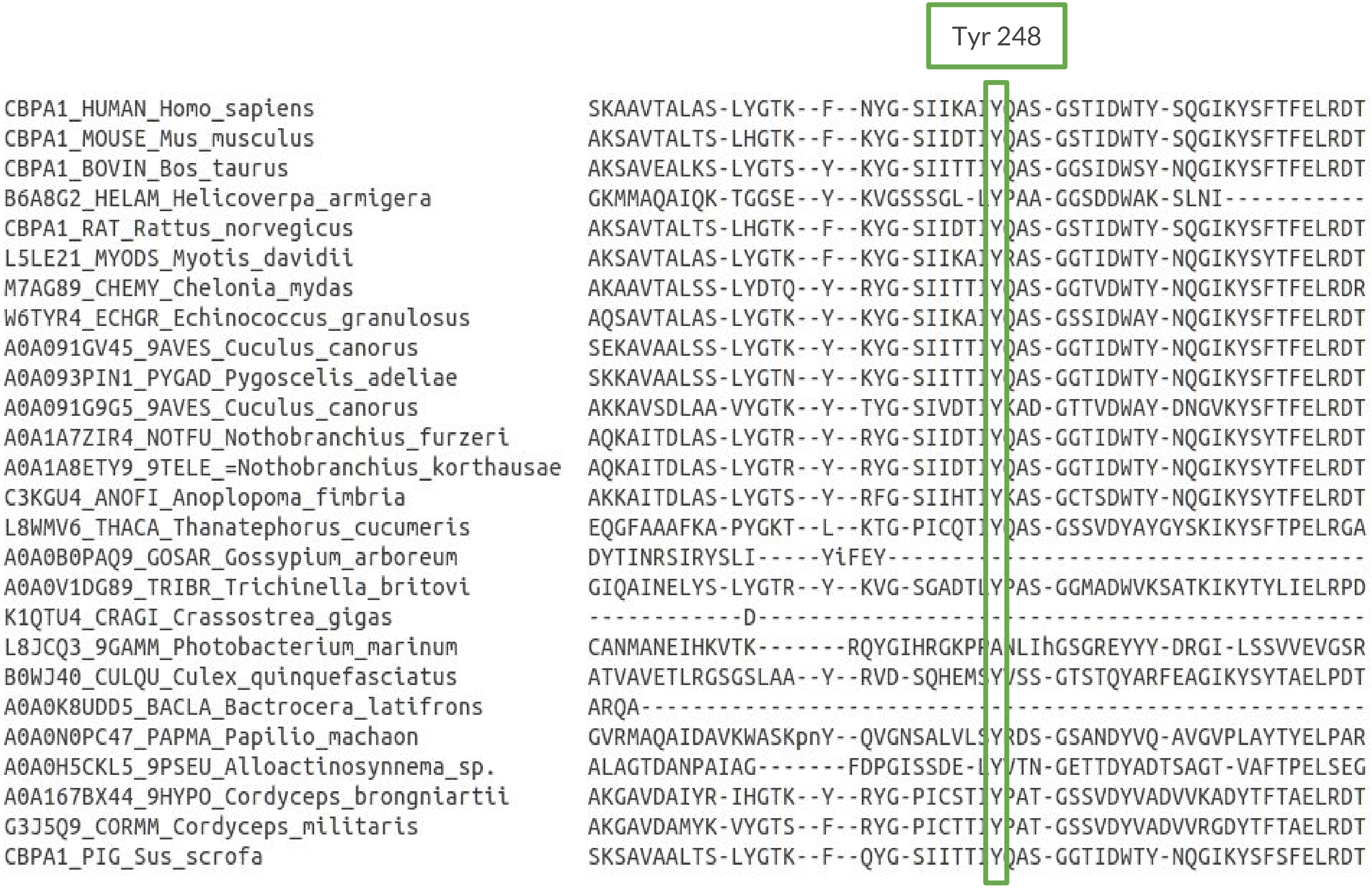


Arg 145

```
CBPA1_HUMAN_Homo_sapiens                    --------------------------lciGVDPNRNWDAGFGLS--GASSNPCSETYHGK
CBPA1_MOUSE_Mus_musculus                    --------------------------lcvGVDPNRNWDAAFGMP--GASSNPCSETYRGK
CBPA1_BOVIN_Bos_taurus                      --------------------------lcvGVDANRNWDAGFGKA--GASSSPCSETYHGK
B6A8G2_HELAM_Helicoverpa_armigera           ------------------------------------------------------------
CBPA1_RAT_Rattus_norvegicus                 --------------------------lcvGVDPNRNWDAGFGMA--GASSNPCSETYRGK
L5LE21_MYODS_Myotis_davidii                 --------------------------scvGVDPNRNWDAGFGKA--GASSNPCSETYHGK
M7AG89_CHEMY_Chelonia_mydas                 --------------------------aciGVDPNRNWDAGFGGS--GASSNPCSETYRGP
W6TYR4_ECHGR_Echinococcus_granulosus        --------------------------icvGVDPNRNWDAGFGLT--GASSNPCSETYHGS
A0A091GV45_9AVES_Cuculus_canorus            --------------------------vcvGVDPNRNWDAGFGGS--GASKNPCSDTYHGP
A0A093PIN1_PYGAD_Pygoscelis_adeliae         --------------------------hcvGVDPNRNWDAGFGGP--GASGNPCSETYHGP
A0A091G9G5_9AVES_Cuculus_canorus            --------------------------lcvGVDPNRNWDAGFGGS--GSSSNPCSETYHGP
A0A1A7ZIR4_NOTFU_Nothobranchius_furzeri     --------------------------scvGVDPNRNWDAGFGGP--GASSSPCSETYRGP
A0A1A8ETY9_9TELE_=Nothobranchius_korthausae --------------------------scvGVDPNRNWDAGFGGP--GASSSPCSETYRGP
C3KGU4_ANOFI_Anoplopoma_fimbria             --------------------------scvGVDPNRNWDAGFGGA--GASGNPCSETFRGP
L8WMV6_THACA_Thanatephorus_cucumeris        --------------------------tcyGRDINRNWAWQWGTG--GSSTSPCAEDYRGA
A0A0B0PAQ9_GOSAR_Gossypium_arboreum         ------------------------------------------------------------
A0A0V1DG89_TRIBR_Trichinella_britovi        ------------------------gqccrGVDLNRNFNFNFGGQ--GSSTDPCDETFQGP
K1QTU4_CRAGI_Crassostrea_gigas              ------------------------------------------------------------
L8JCQ3_9GAMM_Photobacterium_marinum         ----------------------------tfGVDLNRNFGINF----RQSKDTRSNIYGGP
B0WJ40_CULQU_Culex_quinquefasciatus         lnsawmsctplswlqshsslsvntdqhciGTDLNRNWDYRWNEE--GVSRSACSEHYAGY
A0A0K8UDD5_BACLA_Bactrocera_latifrons       ---------------------------cpGVDLNRNFGYKWGGK--GTSANPCAQTYRGS
A0A0N0PC47_PAPMA_Papilio_machaon            -------------------------niclGVDLNRNYDYKWGT---LSSNSPCSDTYHGK
A0A0H5CKL5_9PSEU_Alloactinosynnema_sp.      -------------------------tvgdGVDPNRNFNEHWNYDaeGSSSATSSETYRGA
A0A167BX44_9HYPO_Cordyceps_brongniartii     --------------------------sclGHDINRNWPYKWDVS-gGSSDNPCAEDFRGR
G3J5Q9_CORMM_Cordyceps_militaris            -----------------------------GHDINRNWPYKWDLS-gGASDNPCAEDFRGI
CBPA1_PIG_Sus_scrofa                        --------------------------fcvGVDPNRNWDAGFGGA--GASSNPCSETYHGK
```

Active site

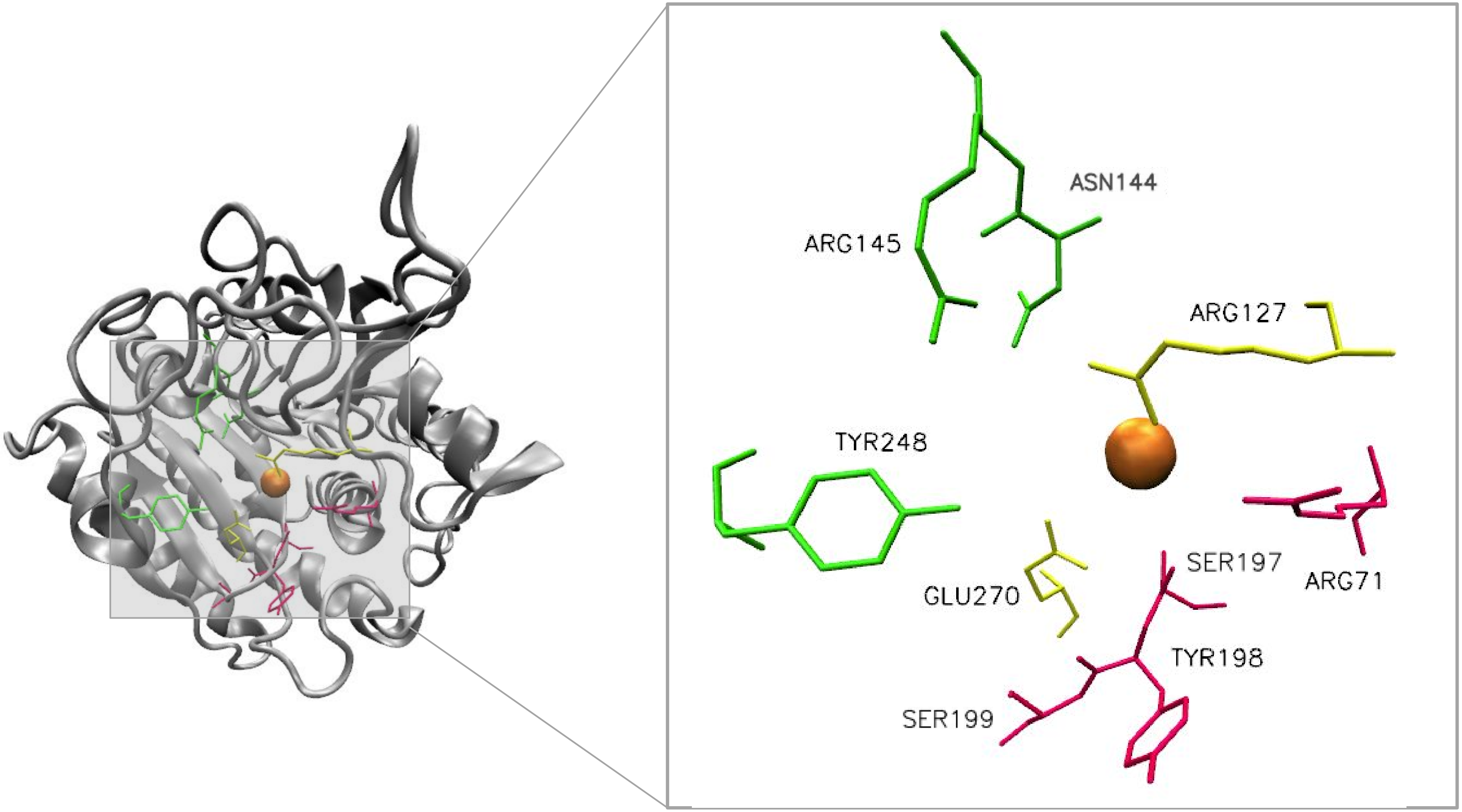
Subsite 1’


Tyr 248

```
CBPA1_HUMAN_Homo_sapiens                      SKAAVTALAS-LYGTK--F--NYG-SIIKAIYQAS-GSTIDWTY-SQGIKYSFTFELRDT
CBPA1_MOUSE_Mus_musculus                      AKSAVTALTS-LHGTK--F--KYG-SIIDTIYQAS-GSTIDWTY-SQGIKYSFTFELRDT
CBPA1_BOVIN_Bos_taurus                        AKSAVEALKS-LYGTS--Y--KYG-SIITTIYQAS-GGSIDWSY-NQGIKYSFTFELRDT
B6A8G2_HELAM_Helicoverpa_armigera             GKMMAQAIQK-TGGSE--Y--KVGSSSGL-LYPAA-GGSDDWAK-SLNI-----------
CBPA1_RAT_Rattus_norvegicus                   AKSAVTALTS-LHGTK--F--KYG-SIIDTIYQAS-GSTIDWTY-SQGIKYSFTFELRDT
L5LE21_MYODS_Myotis_davidii                   AKSAVTALAS-LYGTK--F--KYG-SIIKAIYRAS-GGTIDWTY-NQGIKYSYTFELRDT
M7AG89_CHEMY_Chelonia_mydas                   AKAAVTALSS-LYDTQ--Y--RYG-SIITTIYQAS-GGTVDWTY-NQGIKYSFTFELRDR
W6TYR4_ECHGR_Echinococcus_granulosus          AQSAVTALAS-LYGTK--Y--KYG-SIIKAIYQAS-GSSIDWAY-NQGIKYSFTFELRDT
A0A091GV45_9AVES_Cuculus_canorus              SEKAVAALSS-LYGTK--Y--KYG-SIITTIYQAS-GGTIDWTY-NQGIKYSFTFELRDT
A0A093PIN1_PYGAD_Pygoscelis_adeliae           SKKAVAALSS-LYGTN--Y--KYG-SIITTIYQAS-GGTIDWTY-NQGIKYSFTFELRDT
A0A091G9G5_9AVES_Cuculus_canorus              AKKAVSDLAA-VYGTK--Y--TYG-SIVDTIYKAD-GTTVDWAY-DNGVKYSFTFELRDT
A0A1A7ZIR4_NOTFU_Nothobranchius_furzeri       AQKAITDLAS-LYGTR--Y--RYG-SIIDTIYQAS-GGTIDWTY-NQGIKYSYTFELRDT
A0A1A8ETY9_9TELE_=Nothobranchius_korthausae   AQKAITDLAS-LYGTR--Y--RYG-SIIDTIYQAS-GGTIDWTY-NQGIKYSYTFELRDT
C3KGU4_ANOFI_Anoplopoma_fimbria               AKKAITDLAS-LYGTS--Y--RFG-SIIHTIYKAS-GCTSDWTY-NQGIKYSYTFELRDT
L8WMV6_THACA_Thanatephorus_cucumeris          EQGFAAAFKA-PYGKT--L--KTG-PICQTIYQAS-GSSVDYAYGYSKIKYSFTPELRGA
A0A0B0PAQ9_GOSAR_Gossypium_arboreum           DYTINRSIRYSLI-----YiFEY-------------------------------------
A0A0V1DG89_TRIBR_Trichinella_britovi          GIQAINELYS-LYGTR--Y--KVG-SGADTLYPAS-GGMADWVKSATKIKYTYLIELRPD
K1QTU4_CRAGI_Crassostrea_gigas                ------------D-----------------------------------------------
L8JCQ3_9GAMM_Photobacterium_marinum           CANMANEIHKVTK-------RQYGIHRGKPRANLIhGSGREYYY-DRGI-LSSVVEVGSR
B0WJ40_CULQU_Culex_quinquefasciatus           ATVAVETLRGSGSLAA--Y--RVD-SQHEMSYVSS-GTSTQYARFEAGIKYSYTAELPDT
A0A0K8UDD5_BACLA_Bactrocera_latifrons         ARQA--------------------------------------------------------
A0A0N0PC47_PAPMA_Papilio_machaon              GVRMAQAIDAVKWASKpnY--QVGNSALVLSYRDS-GSANDYVQ-AVGVPLAYTYELPAR
A0A0H5CKL5_9PSEU_Alloactinosynnema_sp.        ALAGTDANPAIAG-------FDPGISSDE-LYVTN-GETTDYADTSAGT-VAFTPELSEG
A0A167BX44_9HYPO_Cordyceps_brongniartii       AKGAVDAIYR-IHGTK--Y--RYG-PICSTIYPAT-GSSVDYVADVVKADYTFTAELRDT
G3J5Q9_CORMM_Cordyceps_militaris              AKGAVDAMYK-VYGTS--F--RYG-PICTTIYPAT-GSSVDYVADVVRGDYTFTAELRDT
CBPA1_PIG_Sus_scrofa                          SKSAVAALTS-LYGTK--F--QYG-SIITTIYQAS-GGTIDWTY-NQGIKYSFSFELRDT
```

Active site

Subsite 2



S1	**Arg 127** **Glu 270**	**S1’**	**Asn 144** **Arg 145** **Tyr 248**	**S2**	**Arg 71** **Ser 197** **Tyr 198** **Ser 199**

Active site

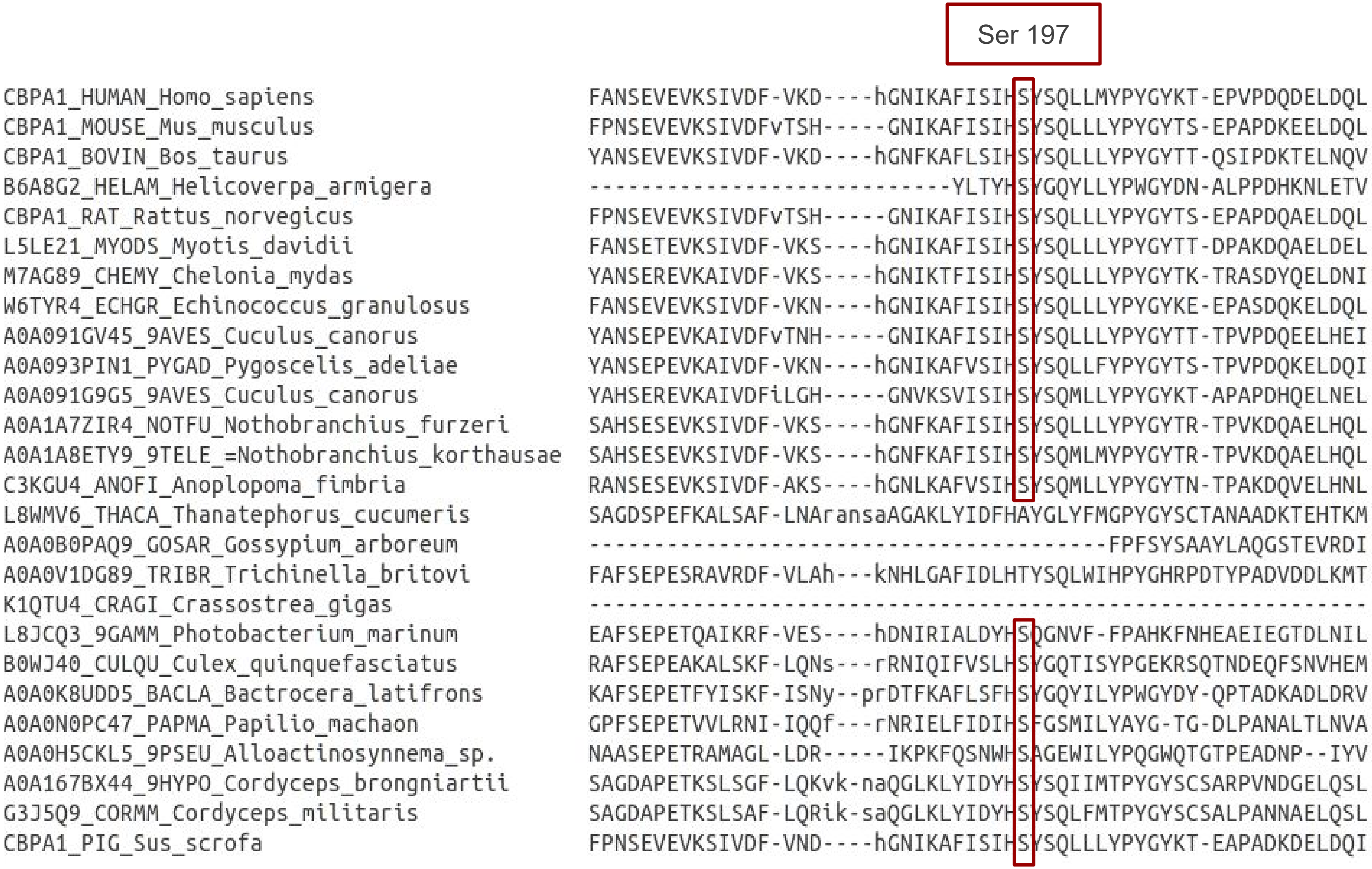
Subsite 2

Arg 71

```
CBPA1_HUMAN_Homo_sapiens                     ----------------GSKRPAIWIDTGIHSREWVTQASGVWFAKKITQDY-gQDAAFTA
CBPA1_MOUSE_Mus_musculus                     ----------------GTNRPAIWIDTGIHSREWVTQASGVWFAKKITKDY-gQEPTLTA
CBPA1_BOVIN_Bos_taurus                       ----------------GSNRPAIWIDLGIHSREWITQATGVWFAKKFTEDY-gQDPSFTA
B6A8G2_HELAM_Helicoverpa_armigera            ------------------------------------------------------------
CBPA1_RAT_Rattus_norvegicus                  ----------------GTNRPAIWIDTGIHSREWVTQASGVWFAKKITKDY-gQDPTFTA
L5LE21_MYODS_Myotis_davidii                  ----------------GNNRPAIWIDTGIHSREWVTQASGIWFAKKITQDY-gQNPAFTA
M7AG89_CHEMY_Chelonia_mydas                  ----------------GTKRPAIWIDTGIHSREWVTQASGVWFAKKIVDSY-gTDPSLTS
W6TYR4_ECHGR_Echinococcus_granulosus         ----------------GNNRPAIWIDTGIHSREWVTQASGVWFAKKITQDY-gHDEVLTS
A0A091GV45_9AVES_Cuculus_canorus             ----------------GTNRPAIWIDTGIHSREWVTQASGVWFAKKIVQDH-eNDSELAS
A0A093PIN1_PYGAD_Pygoscelis_adeliae          ----------------GSNRSAIWLDTGIHSREWITQATGIWTANKIAKEY-gQDPSVTA
A0A091G9G5_9AVES_Cuculus_canorus             ----------------GSNRPAIWLDTGIHSREWITQATGIWTANKIAKEY-gQDPSITA
A0A1A7ZIR4_NOTFU_Nothobranchius_furzeri      ----------------GTNRPAIWIDTGIHSREWVTQASGIWFAKKIVKDY-gSDPALTA
A0A1A8ETY9_9TELE_=Nothobranchius_korthausae  ----------------GTNRPAVWIDTGIHSREWVTQASGIWFAKKIVKDY-gSDPALTA
C3KGU4_ANOFI_Anoplopoma_fimbria              ----------------GTNRPAIWIDTGIHSREWVTQASGTWFAKKIVTDY-gTDPVLTA
L8WMV6_THACA_Thanatephorus_cucumeris         ----------------------------------------EQIAYSLLSNY-tTSSTIKS
A0A0B0PAQ9_GOSAR_Gossypium_arboreum          ------------------------------------------------------------
A0A0V1DG89_TRIBR_Trichinella_britovi         --------------LNASGKKGFWINAGIHAREWASSSTAI--IENCF------------
K1QTU4_CRAGI_Crassostrea_gigas               ------------------------------------------------------------
L8JCQ3_9GAMM_Photobacterium_marinum          --------------AYADLKPALLYTGTIHAREWIGIELAVNFIQHLLDNY-pSNPDVVE
B0WJ40_CULQU_Culex_quinquefasciatus          edkdnvkgkrkkkvTKRGQRSAIFVEAGAHQREWIGPSVATWILDTLTKMVasNDTEL-E
A0A0K8UDD5_BACLA_Bactrocera_latifrons        ------------------------------------------------------------
A0A0N0PC47_PAPMA_Papilio_machaon             --------------FQNNRKPVIVLQSLLHAREWVTLPAALYAIRKLV-----VDITDRD
A0A0H5CKL5_9PSEU_Alloactinosynnema_sp.       -------------gIADGSRPAVLYSSTQHAREWISTEVNRRLLNHYVDRFkaNDPEIKR
A0A167BX44_9HYPO_Cordyceps_brongniartii      --------------SGGGKKPAVVLHGTVHAREWIASMVVEYFINELLSKY-gTDSTITS
G3J5Q9_CORMM_Cordyceps_militaris             --------------SGGGKKPAVVFHGTVHAREWIASMVIEYFINELVTKY-gSDQRITS
CBPA1_PIG_Sus_scrofa                         ----------------GNNRPAIWIDTGIHSREWVTQASGVWFAKKITEDY-gQDPAFTA
```

Active site

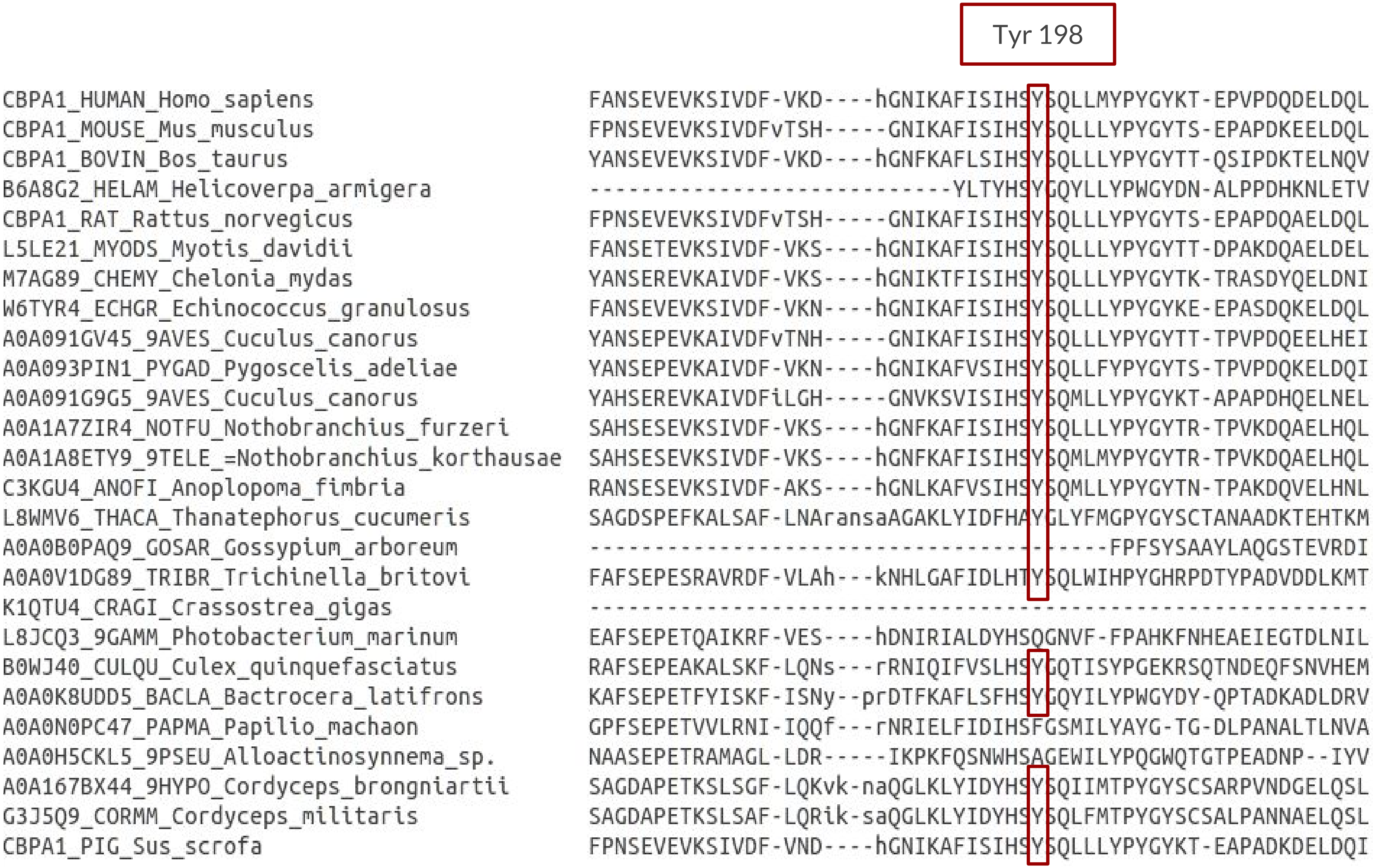
Subsite 2


Ser 197

```
CBPA1_HUMAN_Homo_sapiens                     FANSEVEVKSIVDF-VKD----hGNIKAFISIHSYSQLLMYPYGYKT-EPVPDQDELDQL
CBPA1_MOUSE_Mus_musculus                     FPNSEVEVKSIVDFvTSH-----GNIKAFISIHSYSQLLLYPYGYTS-EPAPDKEELDQL
CBPA1_BOVIN_Bos_taurus                       YANSEVEVKSIVDF-VKD----hGNFKAFLSIHSYSQLLLYPYGYTT-QSIPDKTELNQV
B6A8G2_HELAM_Helicoverpa_armigera            ----------------------------YLTYHSYGQYLLYPWGYDN-ALPPDHKNLETV
CBPA1_RAT_Rattus_norvegicus                  FPNSEVEVKSIVDFvTSH-----GNIKAFISIHSYSQLLLYPYGYTS-EPAPDQAELDQL
L5LE21_MYODS_Myotis_davidii                  FANSETEVKSIVDF-VKS----hGNIKAFISIHSYSQLLLYPYGYTT-DPAKDQAELDEL
M7AG89_CHEMY_Chelonia_mydas                  YANSEREVKAIVDF-VKS----hGNIKTFISIHSYSQLLLYPYGYTK-TRASDYQELDNI
W6TYR4_ECHGR_Echinococcus_granulosus         FANSEVEVKSIVDF-VKN----hGNIKAFISIHSYSQLLLYPYGYKE-EPASDQKELDQL
A0A091GV45_9AVES_Cuculus_canorus             YANSEPEVKAIVDFvTNH-----GNIKAFISIHSYSQLLLYPYGYTT-TPVPDQEELHEI
A0A093PIN1_PYGAD_Pygoscelis_adeliae          YANSEPEVKAIVDF-VKN----hGNIKAFVSIHSYSQLLFYPYGYTS-TPVPDQKELDQI
A0A091G9G5_9AVES_Cuculus_canorus             YAHSEREVKAIVDFiLGH-----GNVKSVISIHSYSQMLLYPYGYKT-APAPDHQELNEL
A0A1A7ZIR4_NOTFU_Nothobranchius_furzeri      SAHSESEVKSIVDF-VKS----hGNFKAFISIHSYSQLLLYPYGYTR-TPVKDQAELHQL
A0A1A8ETY9_9TELE_=Nothobranchius_korthausae  SAHSESEVKSIVDF-VKS----hGNFKAFISIHSYSQMLMYPYGYTR-TPVKDQAELHQL
C3KGU4_ANOFI_Anoplopoma_fimbria              RANSESEVKSIVDF-AKS----hGNLKAFVSIHSYSQMLLYPYGYTN-TPAKDQVELHNL
L8WMV6_THACA_Thanatephorus_cucumeris         SAGDSPEFKALSAF-LNAransaAGAKLYIDFHAYGLYFMGPYGYSCTANAADKTEHTKM
A0A0B0PAQ9_GOSAR_Gossypium_arboreum          ----------------------------------------FPFSYSAAYLAQGSTEVRDI
A0A0V1DG89_TRIBR_Trichinella_britovi         FAFSEPESRAVRDF-VLAh---kNHLGAFIDLHTYSQLWIHPYGHRPDTYPADVDDLKMT
K1QTU4_CRAGI_Crassostrea_gigas               ------------------------------------------------------------
L8JCQ3_9GAMM_Photobacterium_marinum          EAFSEPETQAIKRF-VES----hDNIRIALDYHSQGNVF-FPAHKFNHEAEIEGTDLNIL
B0WJ40_CULQU_Culex_quinquefasciatus          RAFSEPEAKALSKF-LQNs---rRNIQIFVSLHSYGQTISYPGEKRSQTNDEQFSNVHEM
A0A0K8UDD5_BACLA_Bactrocera_latifrons        KAFSEPETFYISKF-ISNy--prDTFKAFLSFHSYGQYILYPWGYDY-QPTADKADLDRV
A0A0N0PC47_PAPMA_Papilio_machaon             GPFSEPETVVLRNI-IQQf---rNRIELFIDIHSFGSMILYAYG-TG-DLPANALTLNVA
A0A0H5CKL5_9PSEU_Alloactinosynnema_sp.       NAASEPETRAMAGL-LDR-----IKPKFQSNWHSAGEWILYPQGWQTGTPEADNP--IYV
A0A167BX44_9HYPO_Cordyceps_brongniartii      SAGDAPETKSLSGF-LQKvk-naQGLKLYIDYHSYSQIIMTPYGYSCSARPVNDGELQSL
G3J5Q9_CORMM_Cordyceps_militaris             SAGDAPETKSLSAF-LQRik-saQGLKLYIDYHSYSQLFMTPYGYSCSALPANNAELQSL
CBPA1_PIG_Sus_scrofa                         FPNSEVEVKSIVDF-VND----hGNIKAFISIHSYSQLLLYPYGYKT-EAPADKDELDQI
```

Active site

Subsite 2

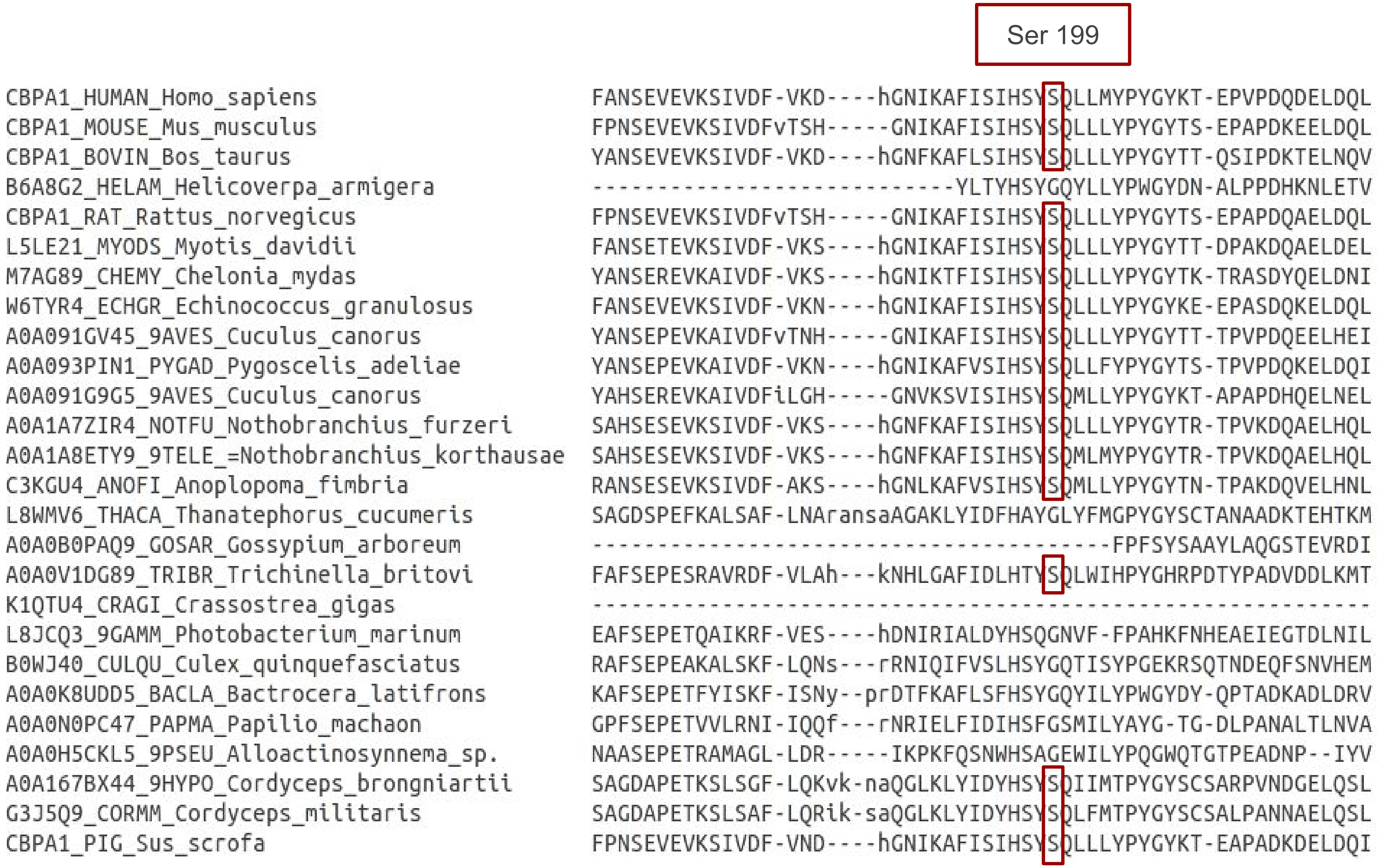


Tyr 198

```
CBPA1_HUMAN_Homo_sapiens                      FANSEVEVKSIVDF-VKD----hGNIKAFISIHSYSQLLMYPYGYKT-EPVPDQDELDQL
CBPA1_MOUSE_Mus_musculus                      FPNSEVEVKSIVDFvTSH-----GNIKAFISIHSYSQLLLYPYGYTS-EPAPDKEELDQL
CBPA1_BOVIN_Bos_taurus                        YANSEVEVKSIVDF-VKD----hGNFKAFLSIHSYSQLLLYPYGYTT-QSIPDKTELNQV
B6A8G2_HELAM_Helicoverpa_armigera             ----------------------------YLTYHSYGQYLLYPWGYDN-ALPPDHKNLETV
CBPA1_RAT_Rattus_norvegicus                   FPNSEVEVKSIVDFvTSH-----GNIKAFISIHSYSQLLLYPYGYTS-EPAPDQAELDQL
L5LE21_MYODS_Myotis_davidii                   FANSETEVKSIVDF-VKS----hGNIKAFISIHSYSQLLLYPYGYTT-DPAKDQAELDEL
M7AG89_CHEMY_Chelonia_mydas                   YANSEREVKAIVDF-VKS----hGNIKTFISIHSYSQLLLYPYGYTK-TRASDYQELDNI
W6TYR4_ECHGR_Echinococcus_granulosus          FANSEVEVKSIVDF-VKN----hGNIKAFISIHSYSQLLLYPYGYKE-EPASDQKELDQL
A0A091GV45_9AVES_Cuculus_canorus              YANSEPEVKAIVDFvTNH-----GNIKAFISIHSYSQLLLYPYGYTT-TPVPDQEELHEI
A0A093PIN1_PYGAD_Pygoscelis_adeliae           YANSEPEVKAIVDF-VKN----hGNIKAFVSIHSYSQLLFYPYGYTS-TPVPDQKELDQI
A0A091G9G5_9AVES_Cuculus_canorus              YAHSEREVKAIVDFiLGH-----GNVKSVISIHSYSQMLLYPYGYKT-APAPDHQELNEL
A0A1A7ZIR4_NOTFU_Nothobranchius_furzeri       SAHSESEVKSIVDF-VKS----hGNFKAFISIHSYSQLLLYPYGYTR-TPVKDQAELHQL
A0A1A8ETY9_9TELE_=Nothobranchius_korthausae   SAHSESEVKSIVDF-VKS----hGNFKAFISIHSYSQMLMYPYGYTR-TPVKDQAELHQL
C3KGU4_ANOFI_Anoplopoma_fimbria               RANSESEVKSIVDF-AKS----hGNLKAFVSIHSYSQMLLYPYGYTN-TPAKDQVELHNL
L8WMV6_THACA_Thanatephorus_cucumeris          SAGDSPEFKALSAF-LNAransaAGAKLYIDFHAYGLYFMGPYGYSCTANAADKTEHTKM
A0A0B0PAQ9_GOSAR_Gossypium_arboreum           ----------------------------------------FPFSYSAAYLAQGSTEVRDI
A0A0V1DG89_TRIBR_Trichinella_britovi          FAFSEPESRAVRDF-VLAh---kNHLGAFIDLHTYSQLWIHPYGHRPDTYPADVDDLKMT
K1QTU4_CRAGI_Crassostrea_gigas                ------------------------------------------------------------
L8JCQ3_9GAMM_Photobacterium_marinum           EAFSEPETQAIKRF-VES----hDNIRIALDYHSQGNVF-FPAHKFNHEAEIEGTDLNIL
B0WJ40_CULQU_Culex_quinquefasciatus           RAFSEPEAKALSKF-LQNs---rRNIQIFVSLHSYGQTISYPGEKRSQTNDEQFSNVHEM
A0A0K8UDD5_BACLA_Bactrocera_latifrons         KAFSEPETFYISKF-ISNy--prDTFKAFLSFHSYGQYILYPWGYDY-QPTADKADLDRV
A0A0N0PC47_PAPMA_Papilio_machaon              GPFSEPETVVLRNI-IQQf---rNRIELFIDIHSFGSMILYAYG-TG-DLPANALTLNVA
A0A0H5CKL5_9PSEU_Alloactinosynnema_sp.        NAASEPETRAMAGL-LDR-----IKPKFQSNWHSAGEWILYPQGWQTGTPEADNP--IYV
A0A167BX44_9HYPO_Cordyceps_brongniartii       SAGDAPETKSLSGF-LQKvk-naQGLKLYIDYHSYSQIIMTPYGYSCSARPVNDGELQSL
G3J5Q9_CORMM_Cordyceps_militaris              SAGDAPETKSLSAF-LQRik-saQGLKLYIDYHSYSQLFMTPYGYSCSALPANNAELQSL
CBPA1_PIG_Sus_scrofa                          FPNSEVEVKSIVDF-VND----hGNIKAFISIHSYSQLLLYPYGYKT-EAPADKDELDQI
```

Active site

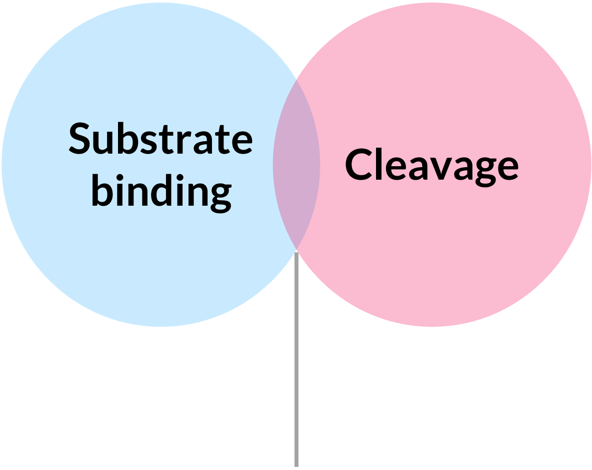
Subsite 2



Ser 199

```
CBPA1_HUMAN_Homo_sapiens                     FANSEVEVKSIVDF-VKD----hGNIKAFISIHSYSQLLMYPYGYKT-EPVPDQDELDQL
CBPA1_MOUSE_Mus_musculus                     FPNSEVEVKSIVDFvTSH-----GNIKAFISIHSYSQLLLYPYGYTS-EPAPDKEELDQL
CBPA1_BOVIN_Bos_taurus                       YANSEVEVKSIVDF-VKD----hGNFKAFLSIHSYSQLLLYPYGYTT-QSIPDKTELNQV
B6A8G2_HELAM_Helicoverpa_armigera            ----------------------------YLTYHSYGQYLLYPWGYDN-ALPPDHKNLETV
CBPA1_RAT_Rattus_norvegicus                  FPNSEVEVKSIVDFvTSH-----GNIKAFISIHSYSQLLLYPYGYTS-EPAPDQAELDQL
L5LE21_MYODS_Myotis_davidii                  FANSETEVKSIVDF-VKS----hGNIKAFISIHSYSQLLLYPYGYTT-DPAKDQAELDEL
M7AG89_CHEMY_Chelonia_mydas                  YANSEREVKAIVDF-VKS----hGNIKTFISIHSYSQLLLYPYGYTK-TRASDYQELDNI
W6TYR4_ECHGR_Echinococcus_granulosus         FANSEVEVKSIVDF-VKN----hGNIKAFISIHSYSQLLLYPYGYKE-EPASDQKELDQL
A0A091GV45_9AVES_Cuculus_canorus             YANSEPEVKAIVDFvTNH-----GNIKAFISIHSYSQLLLYPYGYTT-TPVPDQEELHEI
A0A093PIN1_PYGAD_Pygoscelis_adeliae          YANSEPEVKAIVDF-VKN----hGNIKAFVSIHSYSQLLFYPYGYTS-TPVPDQKELDQI
A0A091G9G5_9AVES_Cuculus_canorus             YAHSEREVKAIVDFiLGH-----GNVKSVISIHSYSQMLLYPYGYKT-APAPDHQELNEL
A0A1A7ZIR4_NOTFU_Nothobranchius_furzeri      SAHSESEVKSIVDF-VKS----hGNFKAFISIHSYSQLLLYPYGYTR-TPVKDQAELHQL
A0A1A8ETY9_9TELE_=Nothobranchius_korthausae  SAHSESEVKSIVDF-VKS----hGNFKAFISIHSYSQMLMYPYGYTR-TPVKDQAELHQL
C3KGU4_ANOFI_Anoplopoma_fimbria              RANSESEVKSIVDF-AKS----hGNLKAFVSIHSYSQMLLYPYGYTN-TPAKDQVELHNL
L8WMV6_THACA_Thanatephorus_cucumeris         SAGDSPEFKALSAF-LNAransaAGAKLYIDFHAYGLYFMGPYGYSCTANAADKTEHTKM
A0A0B0PAQ9_GOSAR_Gossypium_arboreum          ----------------------------------------FPFSYSAAYLAQGSTEVRDI
A0A0V1DG89_TRIBR_Trichinella_britovi         FAFSEPESRAVRDF-VLAh---kNHLGAFIDLHTYSQLWIHPYGHRPDTYPADVDDLKMT
K1QTU4_CRAGI_Crassostrea_gigas               ------------------------------------------------------------
L8JCQ3_9GAMM_Photobacterium_marinum          EAFSEPETQAIKRF-VES----hDNIRIALDYHSQGNVF-FPAHKFNHEAEIEGTDLNIL
B0WJ40_CULQU_Culex_quinquefasciatus          RAFSEPEAKALSKF-LQNs---rRNIQIFVSLHSYGQTISYPGEKRSQTNDEQFSNVHEM
A0A0K8UDD5_BACLA_Bactrocera_latifrons        KAFSEPETFYISKF-ISNy--prDTFKAFLSFHSYGQYILYPWGYDY-QPTADKADLDRV
A0A0N0PC47_PAPMA_Papilio_machaon             GPFSEPETVVLRNI-IQQf---rNRIELFIDIHSFGSMILYAYG-TG-DLPANALTLNVA
A0A0H5CKL5_9PSEU_Alloactinosynnema_sp.       NAASEPETRAMAGL-LDR-----IKPKFQSNWHSAGEWILYPQGWQTGTPEADNP--IYV
A0A167BX44_9HYPO_Cordyceps_brongniartii      SAGDAPETKSLSGF-LQKvk-naQGLKLYIDYHSYSQIIMTPYGYSCSARPVNDGELQSL
G3J5Q9_CORMM_Cordyceps_militaris             SAGDAPETKSLSAF-LQRik-saQGLKLYIDYHSYSQLFMTPYGYSCSALPANNAELQSL
CBPA1_PIG_Sus_scrofa                         FPNSEVEVKSIVDF-VND----hGNIKAFISIHSYSQLLLYPYGYKT-EAPADKDELDQI
```

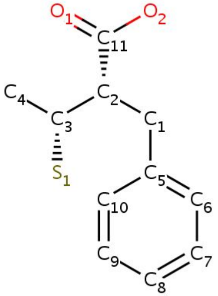
Enzymatic reaction



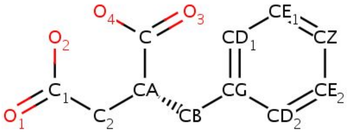
Substrate binding

Cleavage

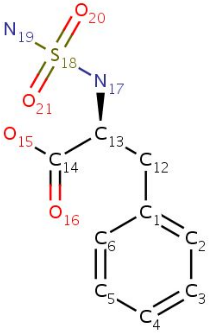
crystallizations with inhibitors



2-benzyl-3,4-epithiobutanoic acid

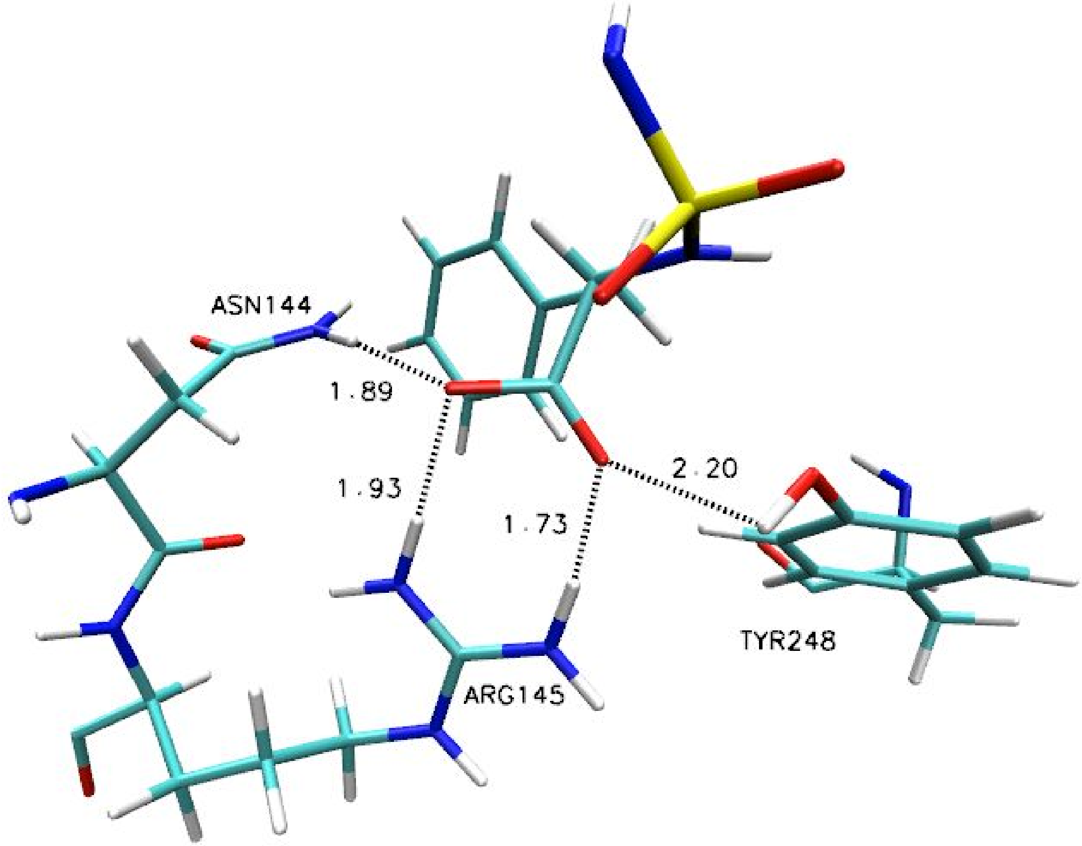


L-benzylsuccinic acid



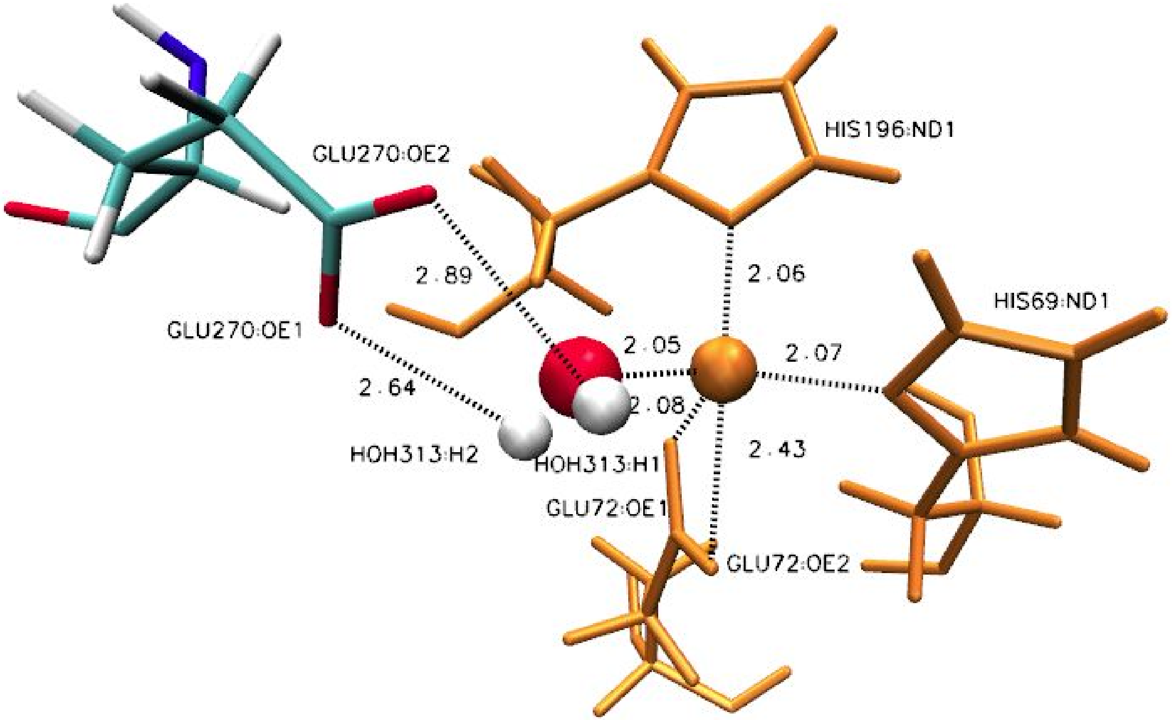
Phenylalanine-N-sulfonamide

Substrate binding



Phenylalanine-N -sulfonamide

Substrate binding

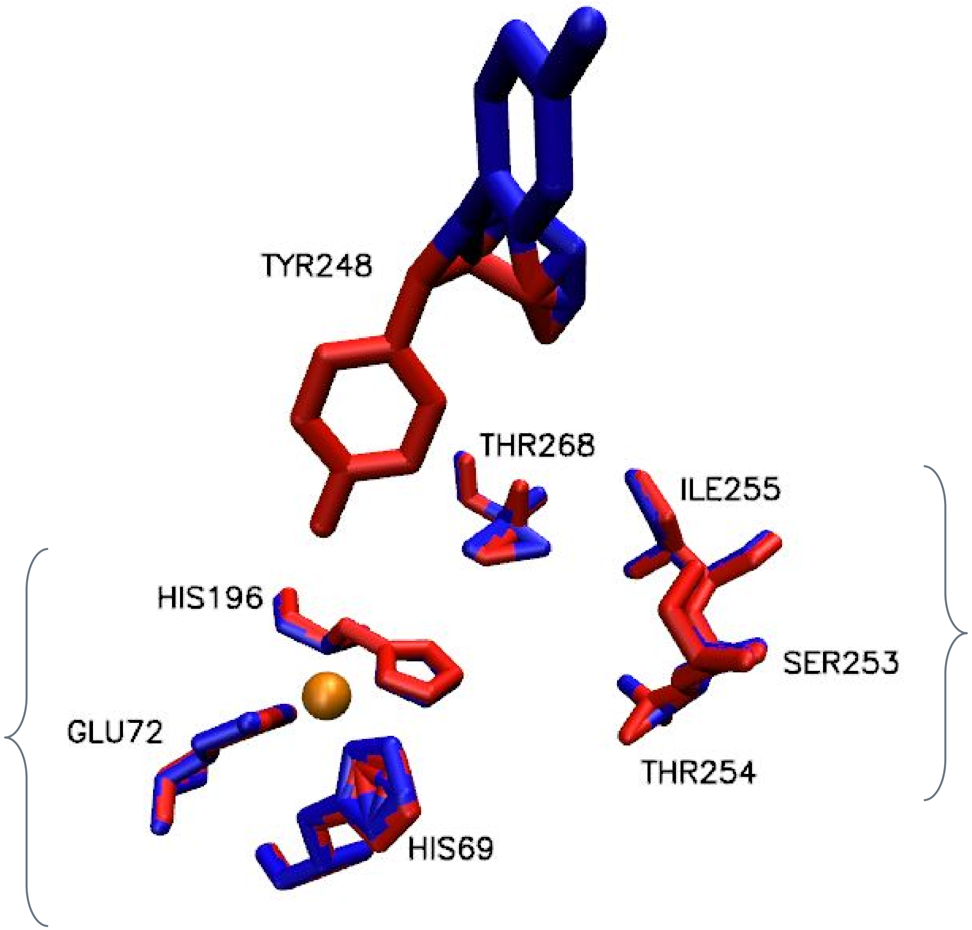


Substrate binding

Conformational changes

🔴 With substrate

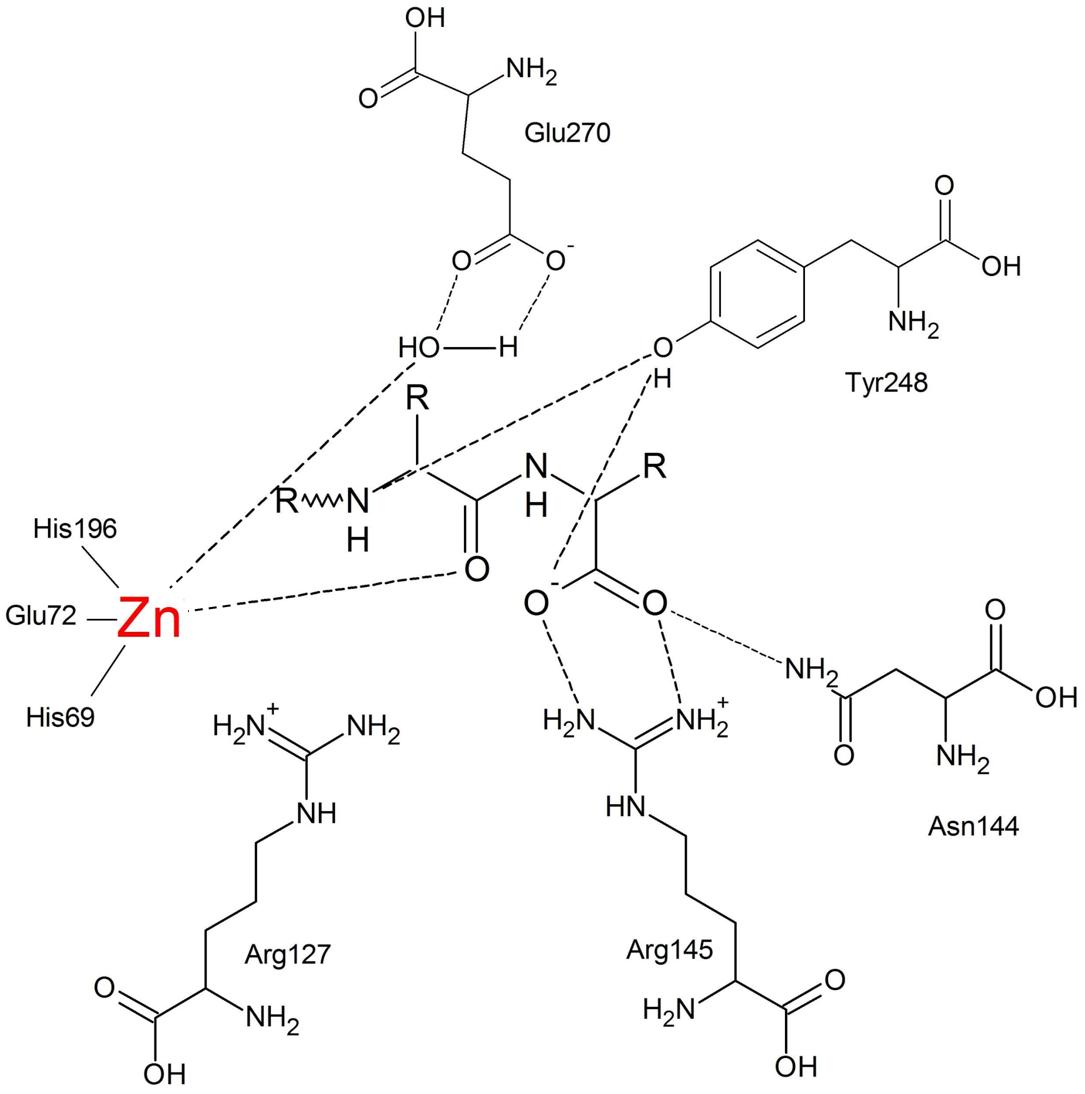
🔵 Without substrate



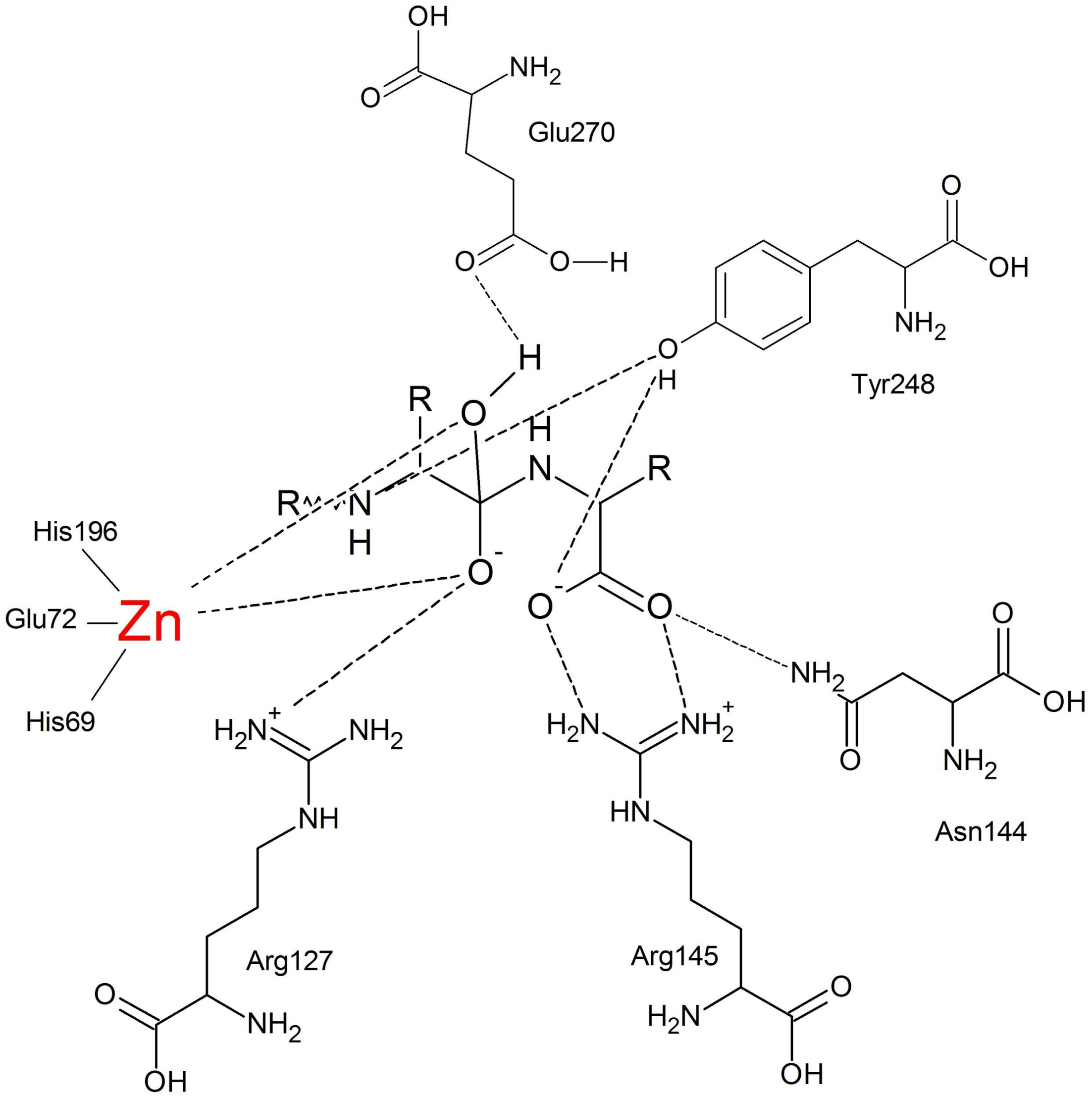
Zn binding residues

Hydrophobic pocket

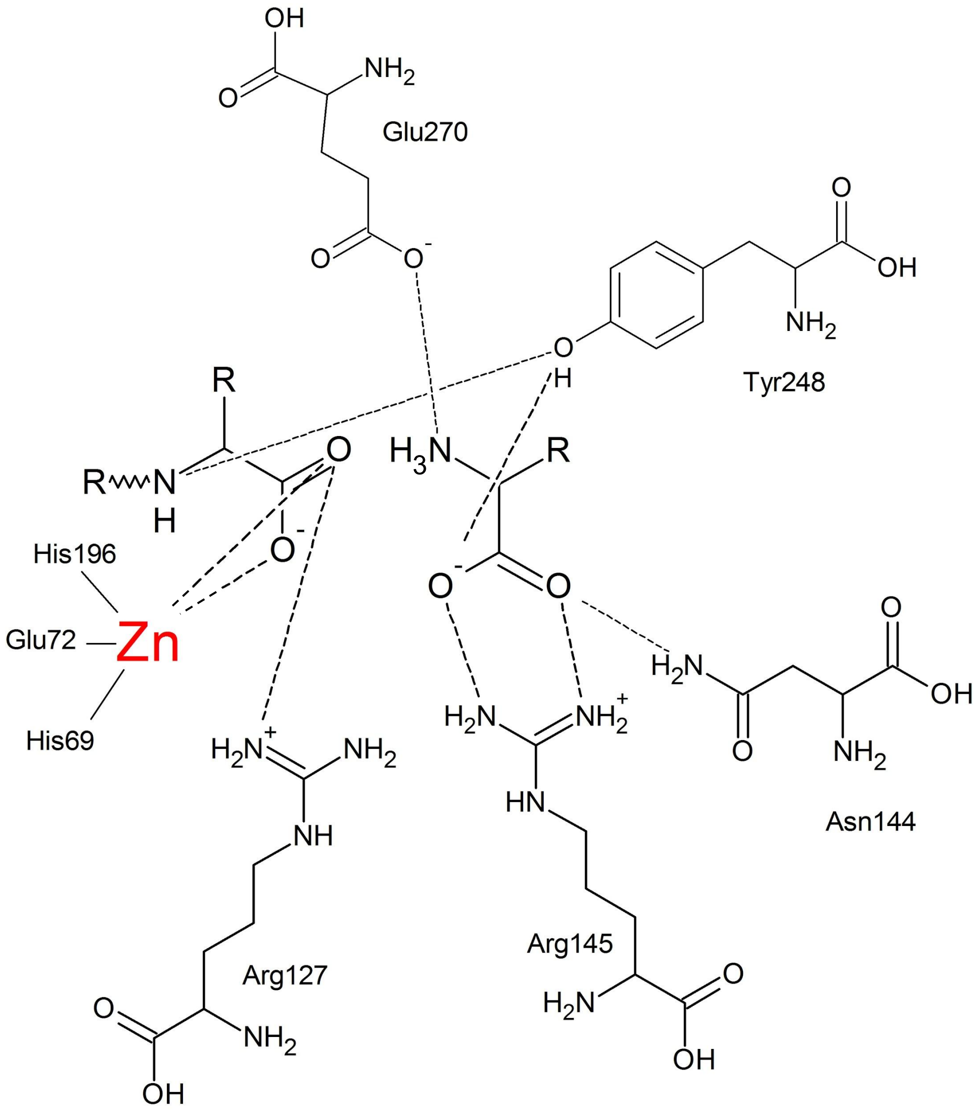
Reaction


OH
NH2
O
Glu270
O
O-
HO
H
O
H
Tyr248
OH
NH2
O
R
R
N
N
H
H
R
O
His196
Glu72
Zn
His69
O-
O
NH2
O
OH
O
NH2
Asn144
H2N+
NH2
NH
Arg127
O
NH2
OH
H2N
NH2+
HN
Arg145
H2N
O
OH

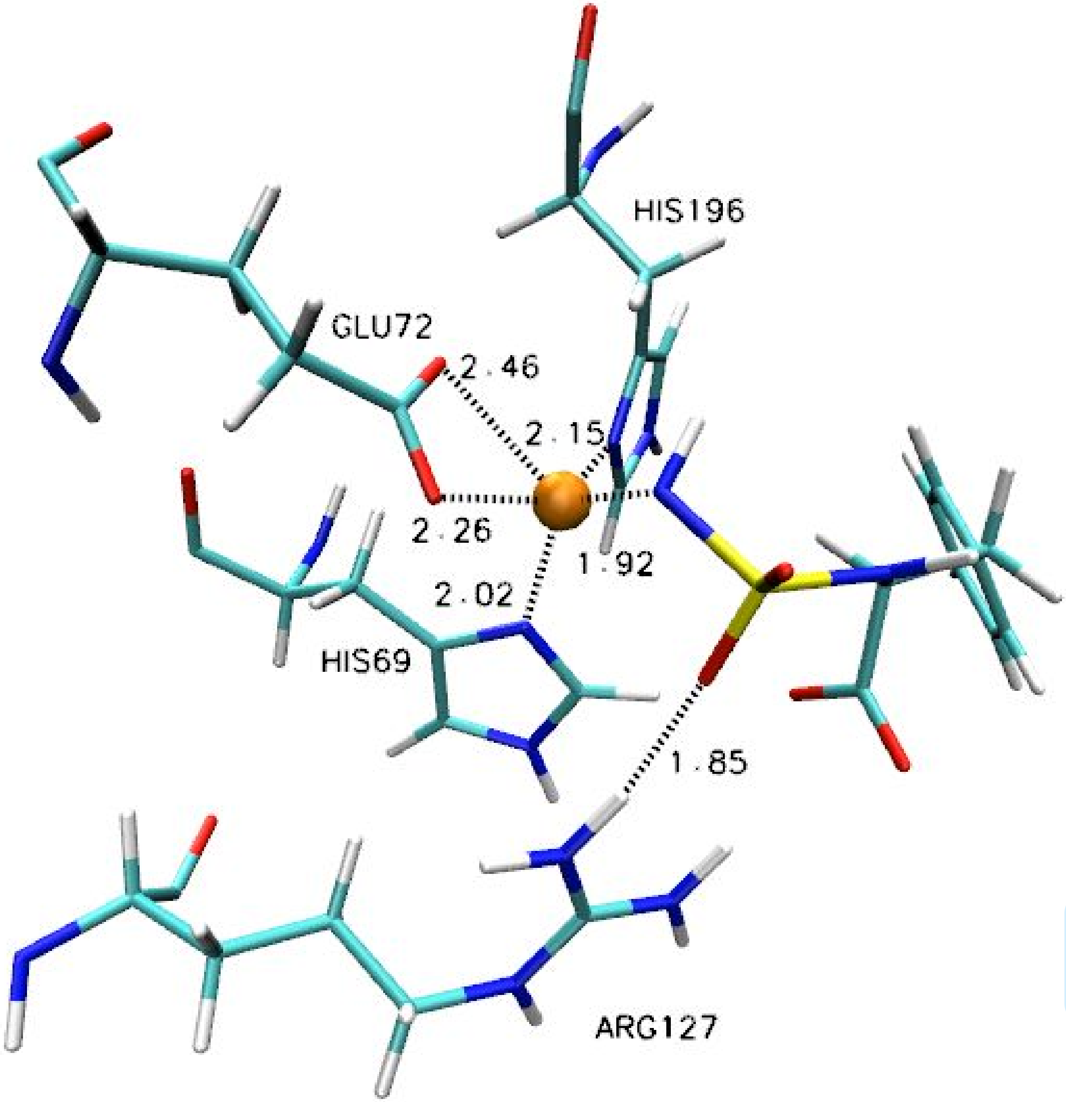
Reaction



Reaction

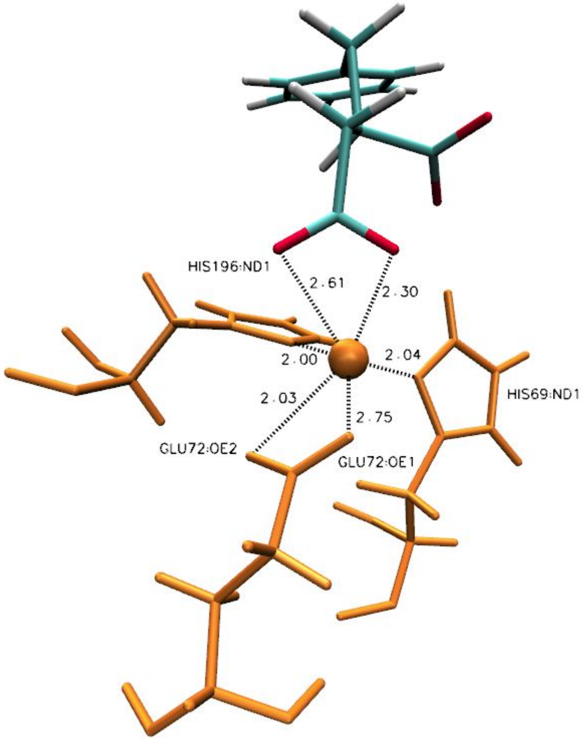


Reaction inhibition

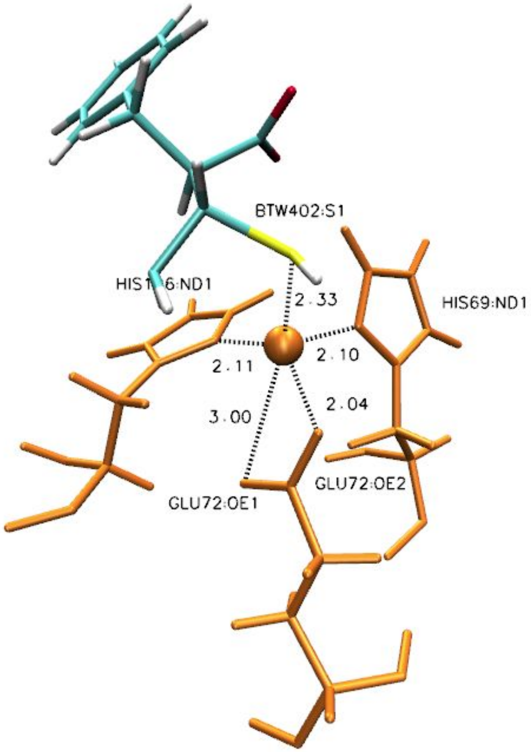


Phenylalanine-N -sulfonamide

Reaction inhibition

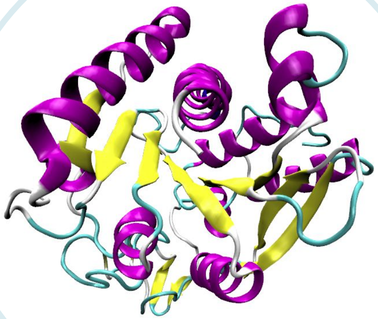


L-benzylsuccinic acid

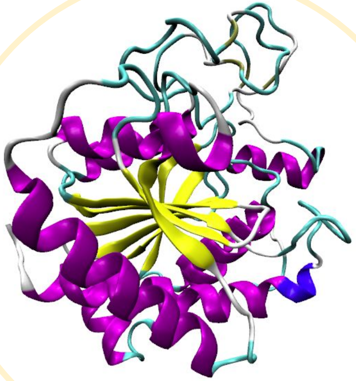


2-benzyl-3,4-epithiobutanoic acid

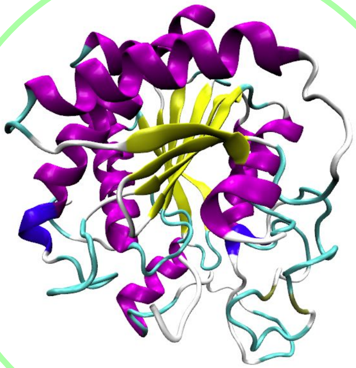
Pancreatic CBP



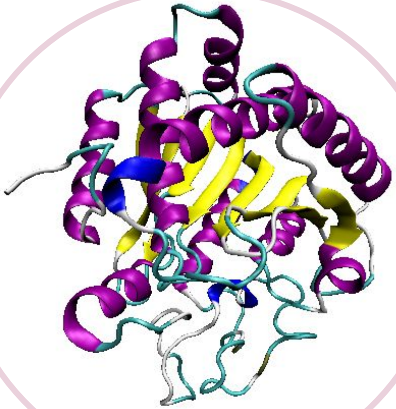
CPA1



CPA2



CPA4



CPB

Pancreatic CBP

Sequence Alignment

Zn binding residues

Subsite 1

Subsite 1’

LEGEND: 2PCU → CPA4 / 4UEE → CPA1 / 1DTD → CPA2 / 1ZLI → CPB

```
2PCU:A|PDBID|CHAIN|SEQUENCE   ..nnfnygayhsleaiy--HEMDNIAADFPDLARRVKIGHSFENRPMYVLKFSTGKGVRR
4UEE:A|PDBID|CHAIN|SEQUENCE   arstdtfnyatyhtleeIYDFLDLLVAENPHLVSKIQIGNTYEGRPIYVLKFSTG-GSKR
1DTD:A|PDBID|CHAIN|SEQUENCE   .....fnfgayhtleei-SQEMDNLVAEHPGLVSKVNIGSSFENRPMNVLKFSTG--GDK
1ZLI:A|PDBID|CHAIN|SEQUENCE   .vratghsyekynnwetIEAWTQQVATENPALISRSVIGTTFEGRAIYLLKVGKA-GQNK

2PCU:A|PDBID|CHAIN|SEQUENCE   PAVWLNAGIHSREWISQATAIWTARKIVSDYqRDPAITSILEKMDIFLLPVANPDGYVYT
4UEE:A|PDBID|CHAIN|SEQUENCE   PAIWIDTGIHSREWVTQASGVWFAKKITQDYgQDAAFTAILDTLDIFLEIVTNPDGFAFT
1DTD:A|PDBID|CHAIN|SEQUENCE   PAIWLDAGIHAREWVTQATALWTANKIVSDYgKDPSITSILDALDIFLLPVTNPDGYVFS
1ZLI:A|PDBID|CHAIN|SEQUENCE   PAIFMDCGFHAREWISPAFCQWFVREAVRTYgREIQVTELLNKLDFYVLPVLNIDGYIYT

2PCU:A|PDBID|CHAIN|SEQUENCE   QTQNRLWRKTRSRNPgs.sciGADPNRNWNASFAGKGASDNPCSEVYHGPHANSEVEVKS
4UEE:A|PDBID|CHAIN|SEQUENCE   HSTNRMWRKTRSHTAgs.lciGVDPNRNWDAGFGLSGASSNPCSETYHGKFANSEVEVKS
1DTD:A|PDBID|CHAIN|SEQUENCE   QTKNRMWRKTRSKVSagslcvGVDPNRNWDAGFGGPGASSNPCSDSYHGPSANSEVEVKS
1ZLI:A|PDBID|CHAIN|SEQUENCE   WTKSRFWRKTRSTHTgs.sciGTDPNRNFDAGWCEIGASRNPCDETYCGPAAESEKETKA

2PCU:A|PDBID|CHAIN|SEQUENCE   VVDFIQK.hGNFKGFIDLHSYSQLLMYPYGYSVKKAPDAEELDKVARLAAKALASVSGTE
4UEE:A|PDBID|CHAIN|SEQUENCE   IVDFVKD.hGNIKAFISIHSYSQLLMYPYGYKTEPVPDQDELDQLSKAAVTALASLYGTK
1DTD:A|PDBID|CHAIN|SEQUENCE   IVDFIKS.hGKVKAFIILHSYSQLLMFPYGYKCTKLDDFDELSEVAQKAAQSLSRLHGTK
1ZLI:A|PDBID|CHAIN|SEQUENCE   LADFIRNklSSIKAYLTIHSYSQMMIYPYSYAYKLGENNAELNALAKATVKELASLHGTK

2PCU:A|PDBID|CHAIN|SEQUENCE   YQVGPTCTTVYPASGSSIDWAY-DNGIKFAFTFELRDTGTygFLLPANQIIPTAEETWLg
4UEE:A|PDBID|CHAIN|SEQUENCE   FNYGSIIKAIYQASGSTIDWTY-SQGIKYSFTFELRDTGRygFLLPASQIIPTAKETWLa
1DTD:A|PDBID|CHAIN|SEQUENCE   YKVGPICSVIYQASGGSIDWSY-DYGIKYSFAFELRDTGRygFLLPARQILPTAEETWLg
1ZLI:A|PDBID|CHAIN|SEQUENCE   YTYGPGATTIYPAAGGSDDWAYDQ-GIRYSFTFELRDTGRygFLLPESQIRATCEETFLa

2PCU:A|PDBID|CHAIN|SEQUENCE   lktimehvrdnl.
4UEE:A|PDBID|CHAIN|SEQUENCE   lltimehtlnhpy
1DTD:A|PDBID|CHAIN|SEQUENCE   lkaimehvrdhpy
1ZLI:A|PDBID|CHAIN|SEQUENCE   ikyvasyvlehly
```

Pancreatic CBP

Sequence Alignment

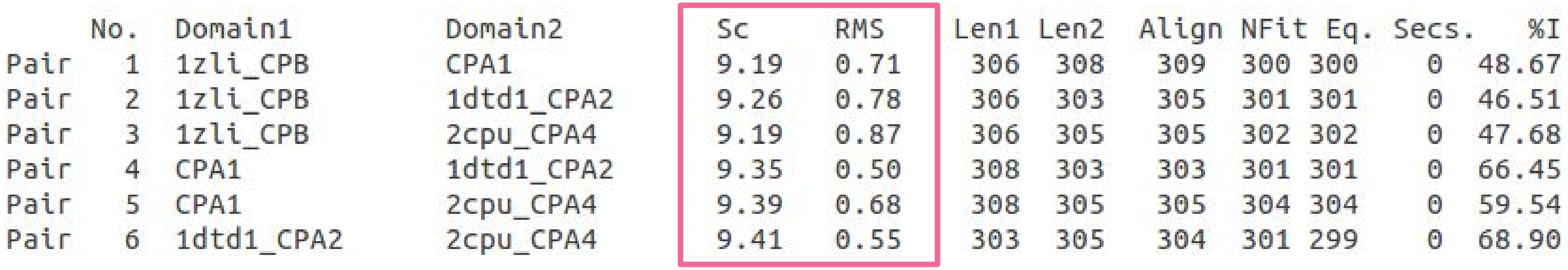
Subsite 2

LEGEND: 2PCU → CPA4 / 4UEE → CPA1 / 1DTD → CPA2 / 1ZLI → CPB

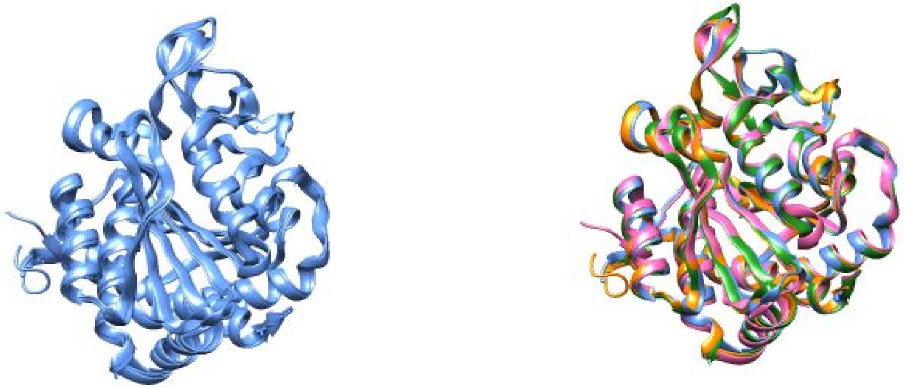
```
2PCU:A|PDBID|CHAIN|SEQUENCE   ..nnfnygayhsleaiy--HEMDNIAADFPDLARRVKIGHSFENRPMYVLKFSTGKGVRR
4UEE:A|PDBID|CHAIN|SEQUENCE   arstdtfnyatyhtleeIYDFLDLLVAENPHLVSKIQIGNTYEGRPIYVLKFSTG-GSKR
1DTD:A|PDBID|CHAIN|SEQUENCE   .....fnfgayhtleei-SQEMDNLVAEHPGLVSKVNIGSSFENRPMNVLKFSTG--GDK
1ZLI:A|PDBID|CHAIN|SEQUENCE   .vratghsyekynnwetIEAWTQQVATENPALISRSVIGTTFEGRAIYLLKVGKA-GQNK

2PCU:A|PDBID|CHAIN|SEQUENCE   PAVWLNAGIHSREWISQATAIWTARKIVSDYqRDPAITSILEKMDIFLLPVANPDGYVYT
4UEE:A|PDBID|CHAIN|SEQUENCE   PAIWIDTGIHSREWVTQASGVWFAKKITQDYgQDAAFTAILDTLDIFLEIVTNPDGFAFT
1DTD:A|PDBID|CHAIN|SEQUENCE   PAIWLDAGIHAREWVTQATALWTANKIVSDYgKDPSITSILDALDIFLLPVTNPDGYVFS
1ZLI:A|PDBID|CHAIN|SEQUENCE   PAIFMDCGFHAREWISPAFCQWFVREAVRTYgREIQVTELLNKLDFYVLPVLNIDGYIYT

2PCU:A|PDBID|CHAIN|SEQUENCE   QTQNRLWRKTRSRNPgs.sciGADPNRNWNASFAGKGASDNPCSEVYHGPHANSEVEVKS
4UEE:A|PDBID|CHAIN|SEQUENCE   HSTNRMWRKTRSHTAgs.lciGVDPNRNWDAGFGLSGASSNPCSETYHGKFANSEVEVKS
1DTD:A|PDBID|CHAIN|SEQUENCE   QTKNRMWRKTRSKVSagslcvGVDPNRNWDAGFGGPGASSNPCSDSYHGPSANSEVEVKS
1ZLI:A|PDBID|CHAIN|SEQUENCE   WTKSRFWRKTRSTHTgs.sciGTDPNRNFDAGWCEIGASRNPCDETYCGPAAESEKETKA

2PCU:A|PDBID|CHAIN|SEQUENCE   VVDFIQK.hGNFKGFIDLHSYSQLLMYPYGYSVKKAPDAEELDKVARLAAKALASVSGTE
4UEE:A|PDBID|CHAIN|SEQUENCE   IVDFVKD.hGNIKAFISIHSYSQLLMYPYGYKTEPVPDQDELDQLSKAAVTALASLYGTK
1DTD:A|PDBID|CHAIN|SEQUENCE   IVDFIKS.hGKVKAFIILHSYSQLLMFPYGYKCTKLDDFDELSEVAQKAAQSLSRLHGTK
1ZLI:A|PDBID|CHAIN|SEQUENCE   LADFIRNklSSIKAYLTIHSYSQMMIYPYSYAYKLGENNAELNALAKATVKELASLHGTK

2PCU:A|PDBID|CHAIN|SEQUENCE   YQVGPTCTTVYPASGSSIDWAY-DNGIKFAFTFELRDTGTygFLLPANQIIPTAEETWLg
4UEE:A|PDBID|CHAIN|SEQUENCE   FNYGSIIKAIYQASGSTIDWTY-SQGIKYSFTFELRDTGRygFLLPASQIIPTAKETWLa
1DTD:A|PDBID|CHAIN|SEQUENCE   YKVGPICSVIYQASGGSIDWSY-DYGIKYSFAFELRDTGRygFLLPARQILPTAEETWLg
1ZLI:A|PDBID|CHAIN|SEQUENCE   YTYGPGATTIYPAAGGSDDWAYDQ-GIRYSFTFELRDTGRygFLLPESQIRATCEETFLa

2PCU:A|PDBID|CHAIN|SEQUENCE   lktimehvrdnl.
4UEE:A|PDBID|CHAIN|SEQUENCE   lltimehtlnhpy
1DTD:A|PDBID|CHAIN|SEQUENCE   lkaimehvrdhpy
1ZLI:A|PDBID|CHAIN|SEQUENCE   ikyvasyvlehly
```

Pancreatic CBP

Superimposition

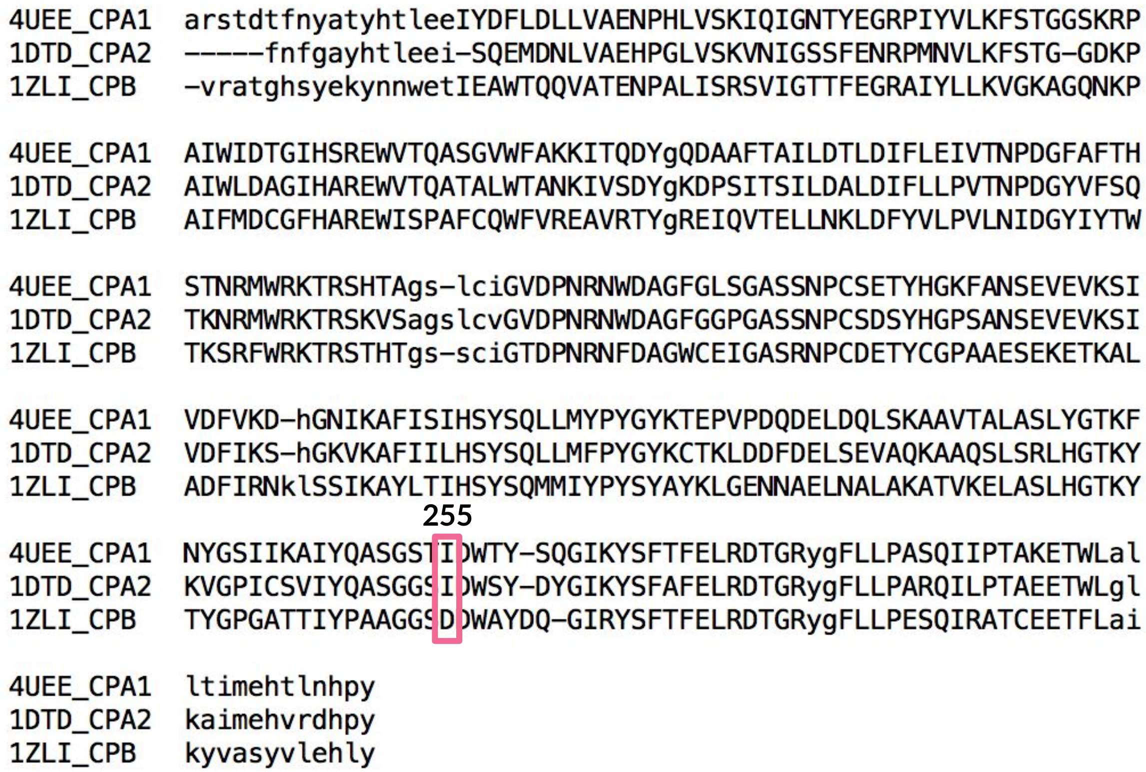


```
      No.  Domain1          Domain2          Sc      RMS     Len1 Len2  Align NFit Eq.  Secs.   %I
Pair   1   1zli_CPB         CPA1             9.19    0.71     306  308   309  300 300      0  48.67
Pair   2   1zli_CPB         1dtd1_CPA2       9.26    0.78     306  303   305  301 301      0  46.51
Pair   3   1zli_CPB         2cpu_CPA4        9.19    0.87     306  305   305  302 302      0  47.68
Pair   4   CPA1             1dtd1_CPA2       9.35    0.50     308  303   303  301 301      0  66.45
Pair   5   CPA1             2cpu_CPA4        9.39    0.68     308  305   305  304 304      0  59.54
Pair   6   1dtd1_CPA2       2cpu_CPA4        9.41    0.55     303  305   304  301 299      0  68.90
```

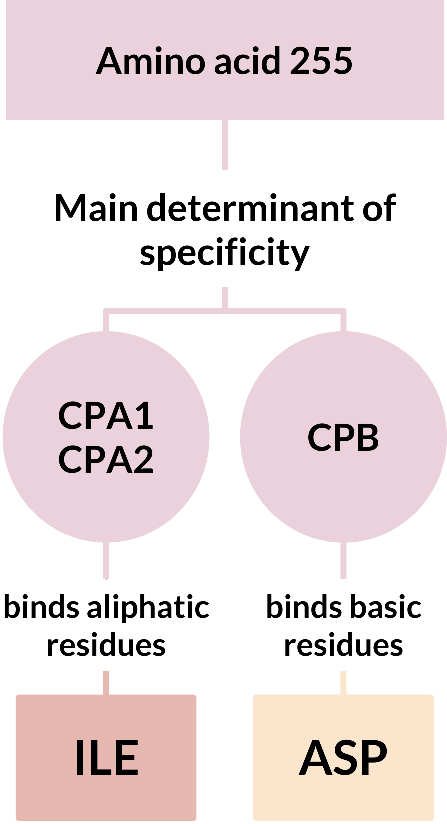


Substrate specificity

A/B subfamily



```
                                       255
4UEE_CPA1  arstdtfnyatyhtleeIYDFLDLLVAENPHLVSKIQIGNTYEGRPIYVLKFSTGGSKRP
1DTD_CPA2  -----fnfgayhtleei-SQEMDNLVAEHPGLVSKVNIGSSFENRPMNVLKFSTG-GDKP
1ZLI_CPB   -vratghsyekynnwetIEAWTQQVATENPALISRSVIGTTFEGRAIYLLKVGKAGQNKP

4UEE_CPA1  AIWIDTGIHSREWVTQASGVWFAKKITQDYgQDAAFTAILDTLDIFLEIVTNPDGFAFTH
1DTD_CPA2  AIWLDAGIHAREWVTQATALWTANKIVSDYgKDPSITSILDALDIFLLPVTNPDGYVFSQ
1ZLI_CPB   AIFMDCGFHAREWISPAFCQWFVREAVRTYgREIQVTELLNKLDFYVLPVLNIDGYIYTW

4UEE_CPA1  STNRMWRKTRSHTAgs-lciGVDPNRNWDAGFGLSGASSNPCSETYHGKFANSEVEVKSI
1DTD_CPA2  TKNRMWRKTRSKVSagslcvGVDPNRNWDAGFGGPGASSNPCSDSYHGPSANSEVEVKSI
1ZLI_CPB   TKSRFWRKTRSTHTgs-sciGTDPNRNFDAGWCEIGASRNPCDETYCGPAAESEKETKAL

4UEE_CPA1  VDFVKD-hGNIKAFISIHSYSQLLMYPYGYKTEPVPDQDELDQLSKAAVTALASLYGTKF
1DTD_CPA2  VDFIKS-hGKVKAFIILHSYSQLLMFPYGYKCTKLDDFDELSEVAQKAAQSLSRLHGTKY
1ZLI_CPB   ADFIRNklSSIKAYLTIHSYSQMMIYPYSYAYKLGENNAELNALAKATVKELASLHGTKY

4UEE_CPA1  NYGSIIKAIYQASGSTIDWTY-SQGIKYSFTFELRDTGRygFLLPASQIIPTAKETWLal
1DTD_CPA2  KVGPICSVIYQASGGSIDWSY-DYGIKYSFAFELRDTGRygFLLPARQILPTAEETWLgl
1ZLI_CPB   TYGPGATTIYPAAGGSDDWAYDQ-GIRYSFTFELRDTGRygFLLPESQIRATCEETFLai

4UEE_CPA1  ltimehtlnhpy
1DTD_CPA2  kaimehvrdhpy
1ZLI_CPB   kyvasyvlehly
```



Amino acid 255

Main determinant of specificity

CPA1
CPA2

CPB

binds aliphatic residues

binds basic residues

ILE

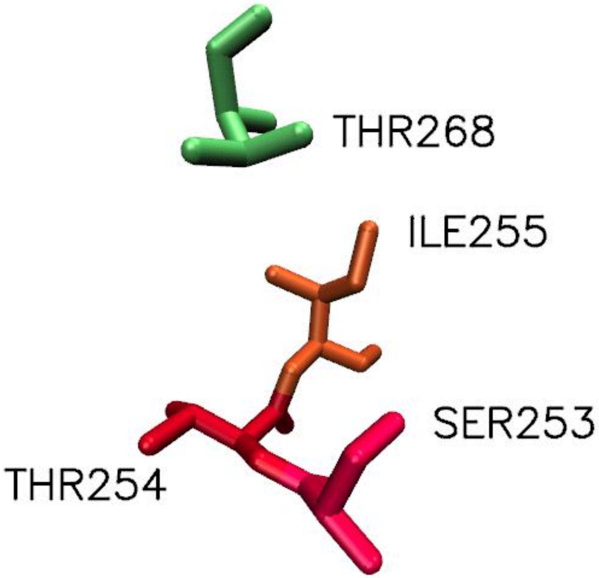
ASP

Substrate specificity A/B subfamily

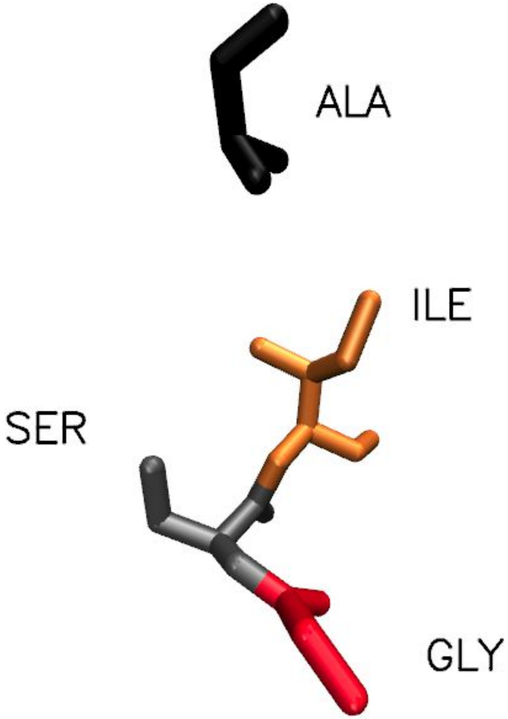
```
111-419-CBPA1    ARSTDTFNYATYHTLEEIYDFLDLLVAENPHLVSKIQIGNTYEGRPIYVL
115-419-CBPA2    ---SGNFNFGAYHTLEEISQEMDNLVAEHPGLVSKVNIGSSFENRPMNVL
                    :..**:.:******* : :* ****:* ****::**.::*.**: **

111-419-CBPA1    KFSTGGSKRPAIWIDTGIHSREWVTQASGVWFAKKITQDYGQDAAFTAIL
115-419-CBPA2    KFSTGGDK-PAIWLDAGIHAREWVTQATALWTANKIVSDYGKDPSITSIL
                 ******.* ****:*:***:*******:.:* *:**..***:*.::*:**

111-419-CBPA1    DTLDIFLEIVTNPDGFAFTHSTNRMWRKTRSHTAGSLCIGVDPNRNWDAG
115-419-CBPA2    DALDIFLLPVTNPDGYVFSQTKNRMWRKTRSKVSGSLCVGVDPNRNWDAG
                 *:*****  ******:.*:::.*********:.:****:***********

111-419-CBPA1    FGLSGASSNPCSETYHGKFANSEVEVKSIVDFVKDHGNIKAFISIHSYSQ
115-419-CBPA2    FGGPGASSNPCSDSYHGPSANSEVEVKSIVDFIKSHGKVKAFITLHSYSQ
                 ** .********::***  *************:*.**::****::*****

111-419-CBPA1    LLMYPYGYKTEPVPDQDELDQLSKAAVTALASLYGTKFNYGSIIKAIYQA
115-419-CBPA2    LLMFPYGYKCTKLDDFDELSEVAQKAAQSLRSLHGTKYKVGPICSVIYQA
                 ***:*****   : * ***.:::: *. :* **:***:: *.* ..****
                   253-255          268
111-419-CBPA1    SGSTIDWTYSQGIKYSFTFELRDTGRYGFLLPASQIIPTAKETWLALLTI
115-419-CBPA2    SGGSIDWSYDYGIKYSFAFELRDTGRYGFLLPARQILPTAEETWLGLKAI
                 **.:***:*. ******:*************** **:***:****.* :*

111-419-CBPA1    MEHTLNHPY
115-419-CBPA2    MEHVRDHPY
                 ***. :***
```

CPA1		CPA2
SER	253	GLY
THR	254	SER
THR	268	ALA

Substrate specificity A/B subfamily

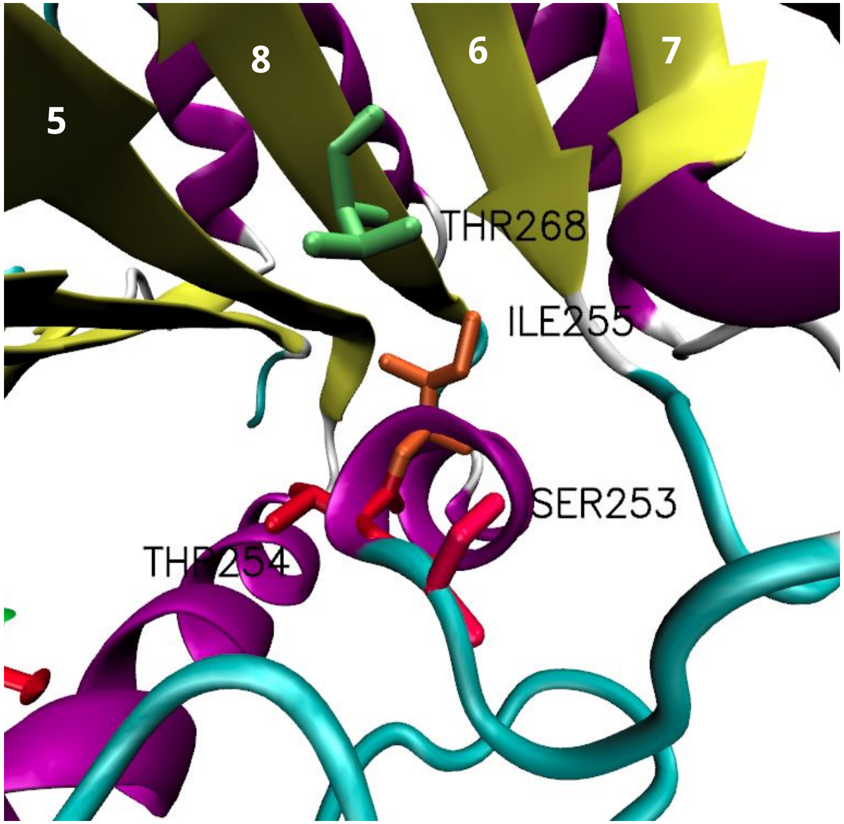


CPA1

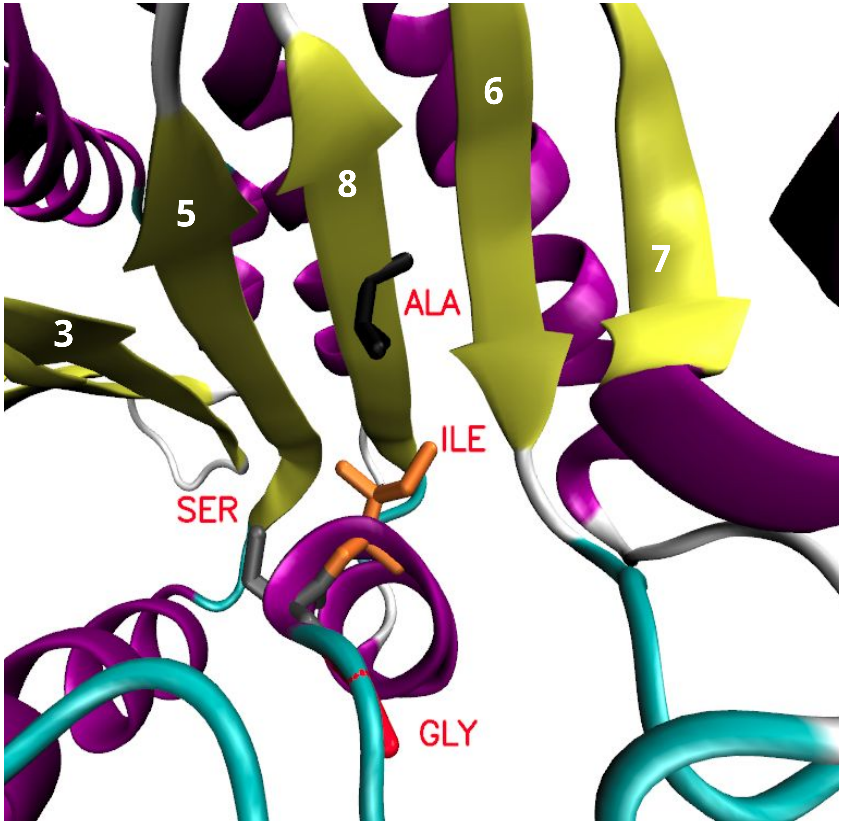


CPA2

Substrate specificity A/B subfamily

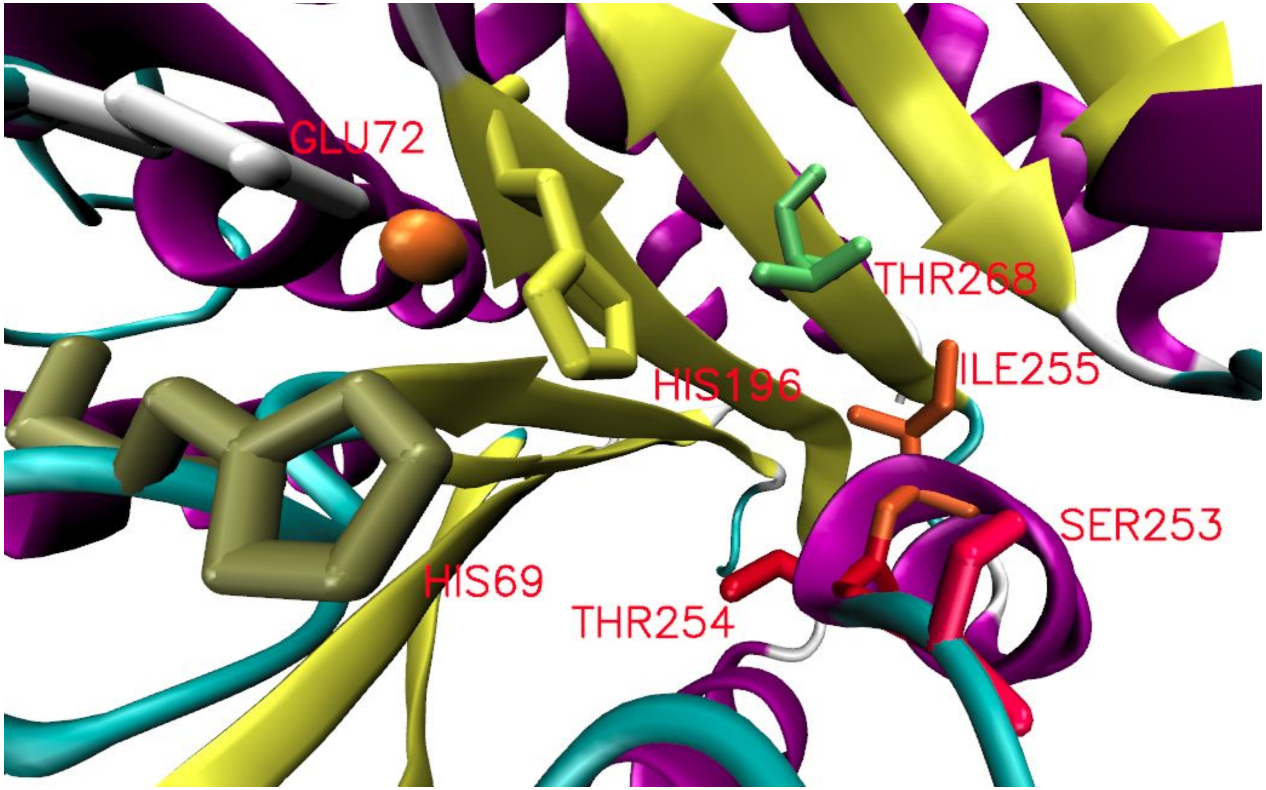


CPA1



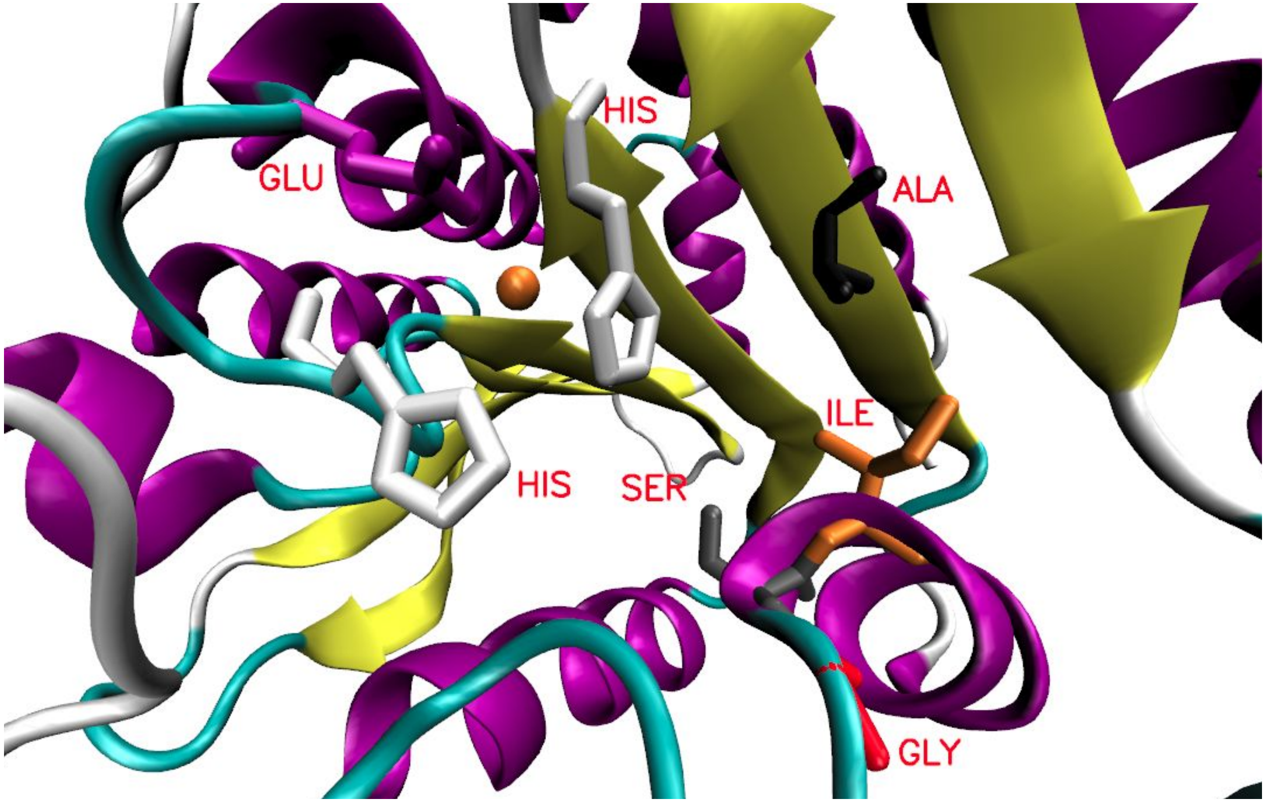
CPA2

Substrate specificity A/B subfamily



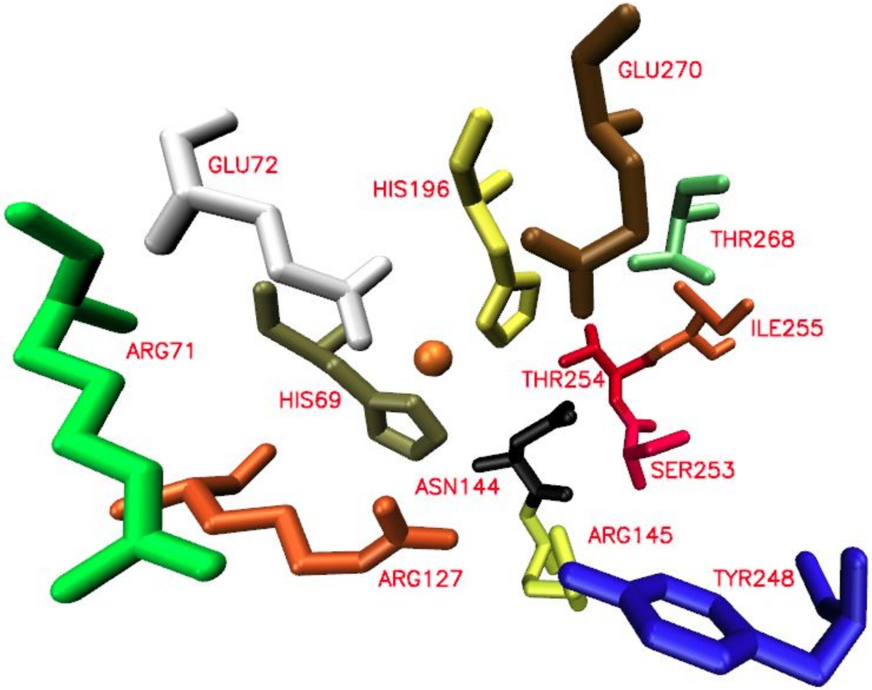
CPA1: specificity pocket + Zn binding residues

Substrate specificity A/B subfamily

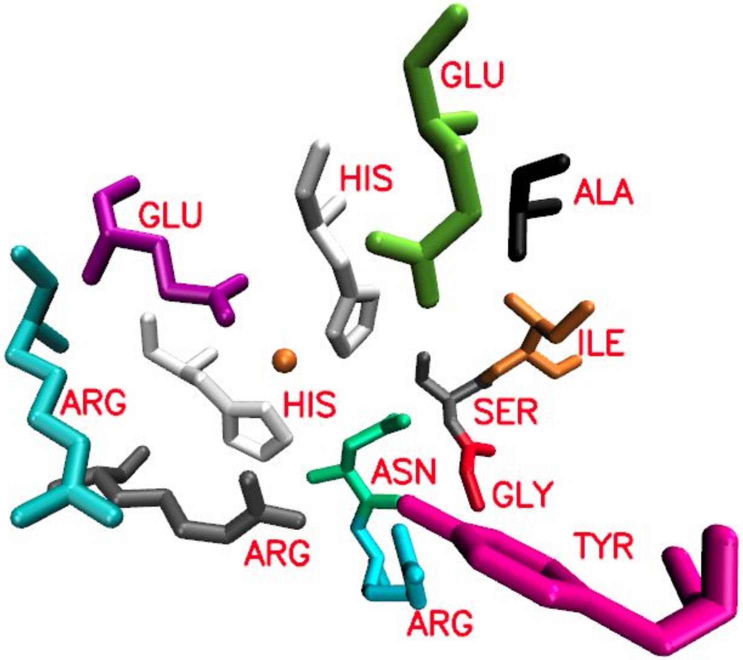


CPA2: specificity pocket + Zn binding residues

Substrate specificity A/B subfamily

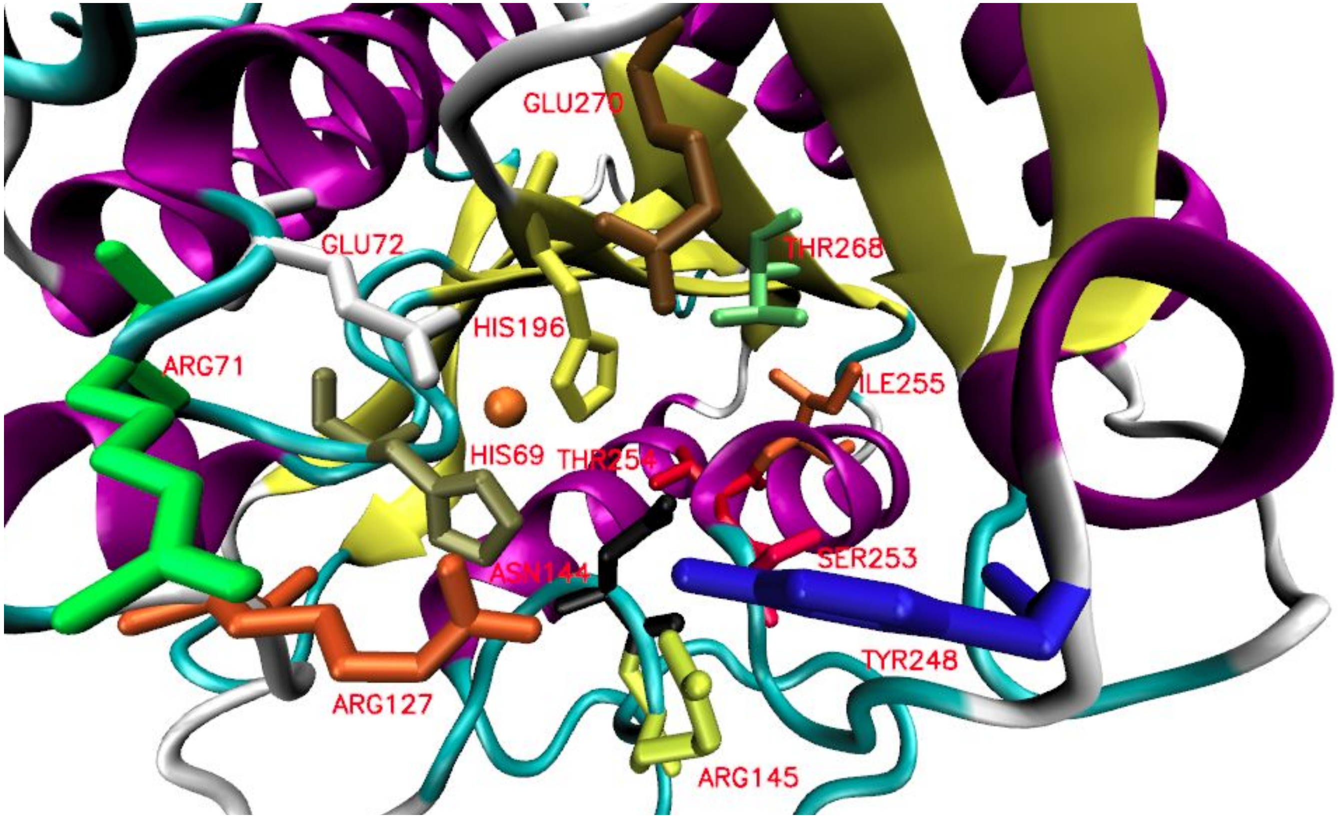


CPA1



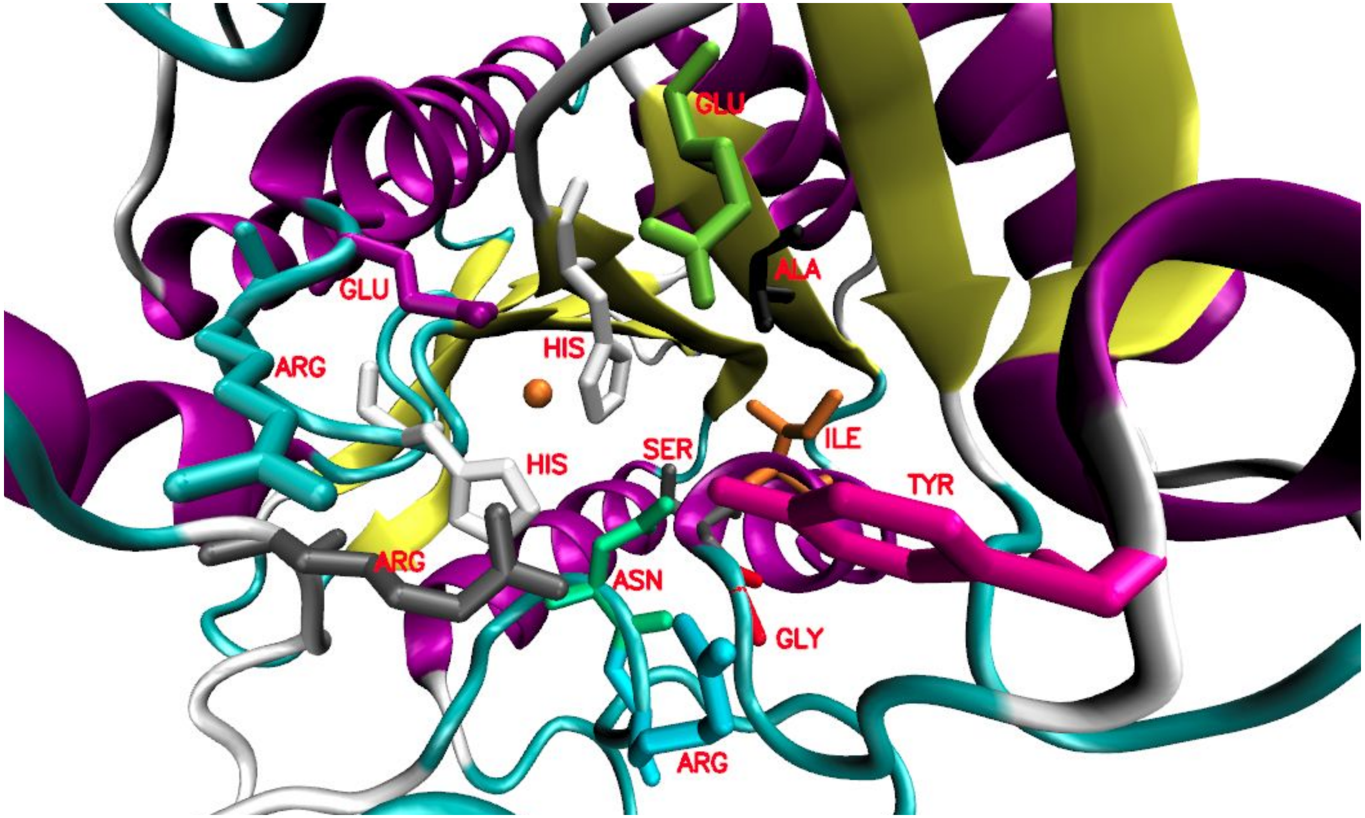
CPA2

Substrate specificity A/B subfamily



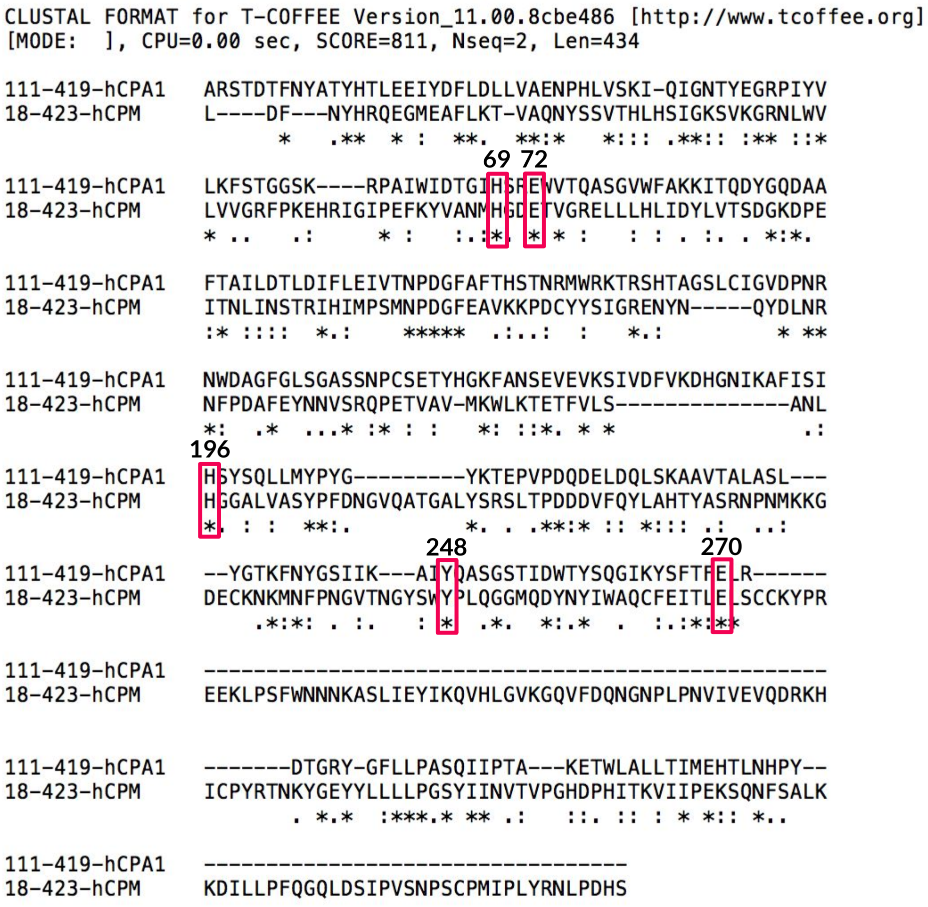
CPA1: specificity pocket + Zn binding residues + active site

Substrate specificity A/B subfamily



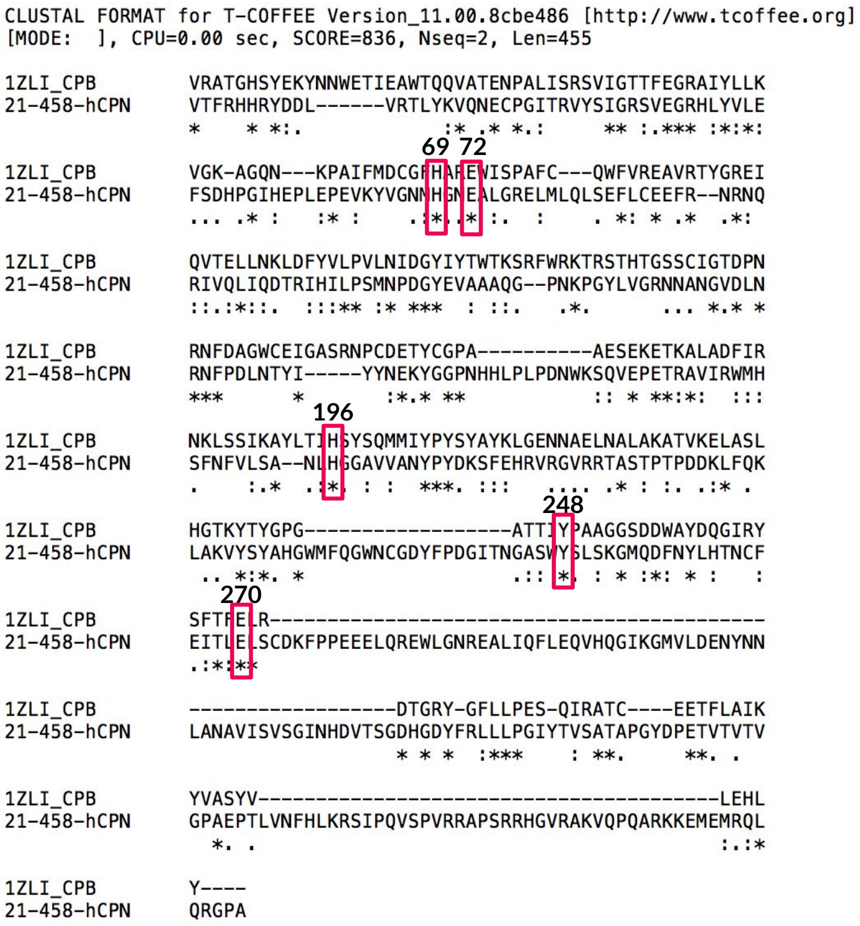
CPA2: specificity pocket + Zn binding residues + active site

Regulatory CBP vs Pancreatic CBP



```
CLUSTAL FORMAT for T-COFFEE Version_11.00.8cbe486 [http://www.tcoffee.org]
[MODE:  ], CPU=0.00 sec, SCORE=811, Nseq=2, Len=434

111-419-hCPA1   ARSTDTFNYATYHTLEEIYDFLDLLVAENPHLVSKI-QIGNTYEGRPIYV
18-423-hCPM     L----DF---NYHRQEGMEAFLKT-VAQNYSSVTHLHSIGKSVKGRNLWV
                      *   .**  * :  **.  **:*   *::: .**:: :** ::*

111-419-hCPA1   LKFSTGGSK----RPAIWIDTGIHSREWVTQASGVWFAKKITQDYGQDAA
18-423-hCPM     LVVGRFPKEHRIGIPEFKYVANMHGDETVGRELLLHLIDYLVTSDGKDPE
                * ..   .:     * :   :.:*. * * :   : : . :. . *:*.

111-419-hCPA1   FTAILDTLDIFLEIVTNPDGFAFTHSTNRMWRKTRSHTAGSLCIGVDPNR
18-423-hCPM     ITNLINSTRIHIMPSMNPDGFEAVKKPDCYYSIGRENYN-----QYDLNR
                :* ::::  *.:    *****  .:..:  :   *.:         * **

111-419-hCPA1   NWDAGFGLSGASSNPCSETYHGKFANSEVEVKSIVDFVKDHGNIKAFISI
18-423-hCPM     NFPDAFEYNNVSRQPETVAV-MKWLKTETFVLS--------------ANL
                *:  .*  ...* :* : :   *: ::*. * *               .:

111-419-hCPA1   HSYSQLLMYPYG---------YKTEPVPDQDELDQLSKAAVTALASL---
18-423-hCPM     HGGALVASYPFDNGVQATGALYSRSLTPDDDVFQYLAHTYASRNPNMKKG
                *. : :  **:.         *. . .**:* :: *::: .:  ..:

111-419-hCPA1   --YGTKFNYGSIIK---AIYQASGSTIDWTYSQGIKYSFTFELR------
18-423-hCPM     DECKNKMNFPNGVTNGYSWYPLQGGMQDYNYIWAQCFEITLELSCCKYPR
                    .*:*: . :.   : *  .*.  *:.*  .  :.:*:**

111-419-hCPA1   --------------------------------------------------
18-423-hCPM     EEKLPSFWNNNKASLIEYIKQVHLGVKGQVFDQNGNPLPNVIVEVQDRKH

111-419-hCPA1   -------DTGRY-GFLLPASQIIPTA---KETWLALLTIMEHTLNHPY--
18-423-hCPM     ICPYRTNKYGEYYLLLLPGSYIINVTVPGHDPHITKVIIPEKSQNFSALK
                       . *.*  :***.* ** .:   ::. :: : * *:: *..

111-419-hCPA1   -----------------------------------
18-423-hCPM     KDILLPFQGQLDSIPVSNPSCPMIPLYRNLPDHS
```

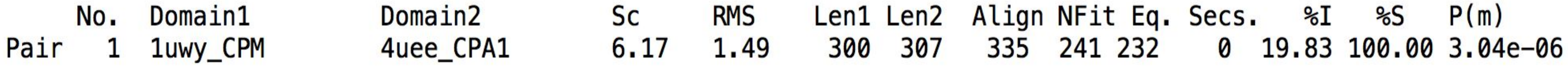
%ID: 15,9



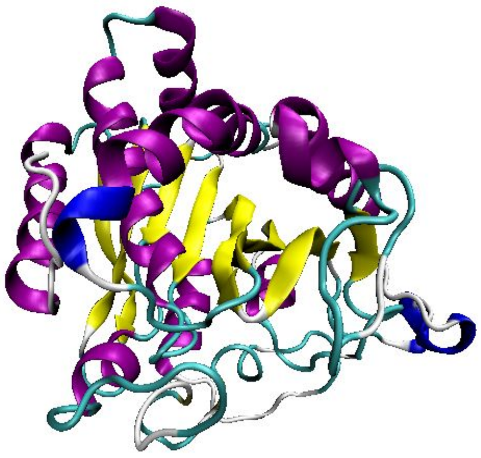
```
CLUSTAL FORMAT for T-COFFEE Version_11.00.8cbe486 [http://www.tcoffee.org]
[MODE:  ], CPU=0.00 sec, SCORE=836, Nseq=2, Len=455

1ZLI_CPB        VRATGHSYEKYNNWETIEAWTQQVATENPALISRSVIGTTFEGRAIYLLK
21-458-hCPN     VTFRHHRYDDL------VRTLYKVQNECPGITRVYSIGRSVEGRHLYVLE
                *    * *:.            :* .* *.:     ** :.*** :*:*:

1ZLI_CPB        VGK-AGQN---KPAIFMDCGFHAREWISPAFC---QWFVREAVRTYGREI
21-458-hCPN     FSDHPGIHEPLEPEVKYVGNMHGNEALGRELMLQLSEFLCEEFR--NRNQ
                ... .* :   :* :    .:*..* :.  :    . *: * .*  .*:

1ZLI_CPB        QVTELLNKLDFYVLPVLNIDGYIYTWTKSRFWRKTRSTHTGSSCIGTDPN
21-458-hCPN     RIVQLIQDTRIHILPSMNPDGYEVAAAQG--PNKPGYLVGRNNANGVDLN
                ::.:*::.  :::** :* ***  : ::.   .*.      ... *.* *

1ZLI_CPB        RNFDAGWCEIGASRNPCDETYCGPA----------AESEKETKALADFIR
21-458-hCPN     RNFPDLNTYI-----YYNEKYGGPNHHLPLPDNWKSQVEPETRAVIRWMH
                ***      *       :*.* **           :: * **:*:  :::

1ZLI_CPB        NKLSSIKAYLTIHSYSQMMIYPYSYAYKLGENNAELNALAKATVKELASL
21-458-hCPN     SFNFVLSA--NLHGGAVVANYPYDKSFEHRVRGVRRTASTPTPDDKLFQK
                .    :.*  .:*. : :  ***. :::   .... .* : :. .:* .

1ZLI_CPB        HGTKYTYGPG------------------ATTIYPAAGGSDDWAYDQGIRY
21-458-hCPN     LAKVYSYAHGWMFQGWNCGDYFPDGITNGASWYSLSKGMQDFNYLHTNCF
                 .. *:*. *                  .:: *. : * :*: * :   :

1ZLI_CPB        SFTFELR-------------------------------------------
21-458-hCPN     EITLELSCDKFPPEEELQREWLGNREALIQFLEQVHQGIKGMVLDENYNN
                .:*:**

1ZLI_CPB        ------------------DTGRY-GFLLPES-QIRATC----EETFLAIK
21-458-hCPN     LANAVISVSGINHDVTSGDHGDYFRLLLPGIYTVSATAPGYDPETVTVTV
                                  * * *  :***    : **.     **. .

1ZLI_CPB        YVASYV--------------------------------------LEHL
21-458-hCPN     GPAEPTLVNFHLKRSIPQVSPVRRAPSRRHGVRAKVQPQARKKEMEMRQL
                  *. .                                        :.:*

1ZLI_CPB        Y----
21-458-hCPN     QRGPA
```

%ID: 16,04

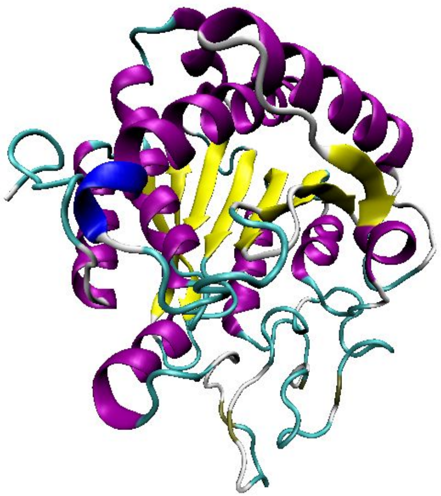
Regulatory CBP vs Pancreatic CBP



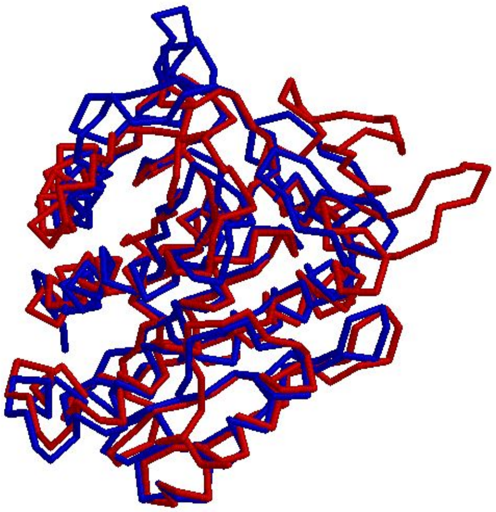
```
      No.  Domain1          Domain2          Sc      RMS     Len1 Len2  Align NFit Eq. Secs.    %I     %S    P(m)
Pair    1  1uwy_CPM         4uee_CPA1        6.17    1.49     300  307   335  241 232     0  19.83 100.00 3.04e-06
```



CPM



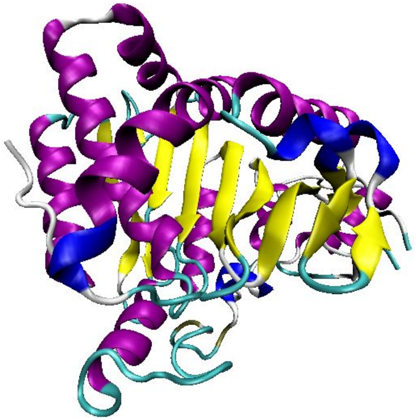
CPA1



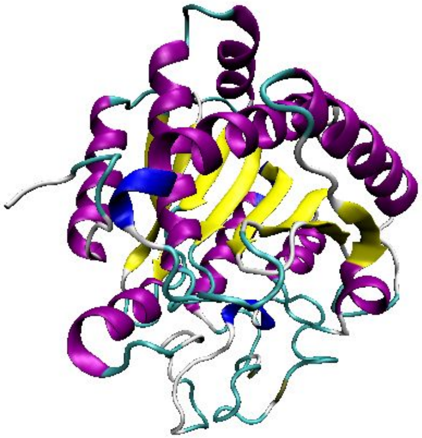
CPM vs CPA1

Regulatory CBP vs Pancreatic CBP

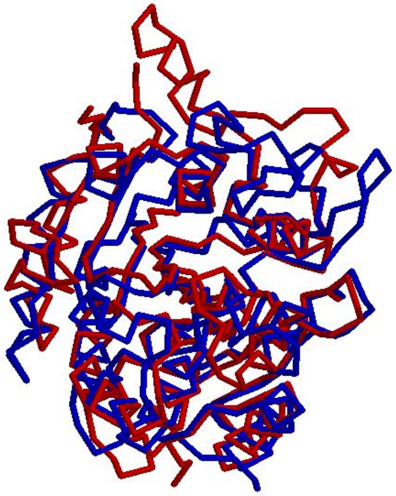
```
     No.  Domain1         Domain2         Sc      RMS     Len1 Len2  Align NFit Eq.  Secs.    %I      %S    P(m)
Pair   1  2nsm_hCPN       1zli_CPB        5.88    1.34     310  306    348  236 232      0  21.98 100.00 3.85e-08
```



CPN



CPB



CPN vs CPB

Conclusions

Overall sequence is poorly maintained among M14 carboxypeptidase subfamilies

Important functional residues are highly conserved along evolution in CPA1 and maintained among M14 carboxypeptidases

Structure is mostly maintained in M14 carboxypeptidases

Functional differences between M14 carboxypeptidases are correlated with changes in sequence and active site properties

Bibliography

Articles

Avilés FX, Vendrell J, Guasch A, Coll M, Huber R. Advances in metallo-procarboxypeptidases. Emerging details on the inhibition mechanism and on the activation process. Eur J Biochem. 1993;211(3):381-389.

Catasus L, Vendrell J, Avilés FX, Carreira S, Puigserver A, Billeter M. The sequence and conformation of human pancreatic procarboxypeptidase A2. cDNA cloning, sequence analysis, and three-dimensional model. J Biol Chem. 1995;270(12):6651-7.

Fernández D, Boix E, Pallarès I, Avilés F, Vendrell J. Structural and Functional Analysis of the Complex between Citrate and the Zinc Peptidase Carboxypeptidase A. Enzyme Res. 2011;2011:1-8.

Fernández D, Pallares I, Covaleda G, Aviles FX, Vendrell J. Metallocarboxypeptidases and their inhibitors: recent developments in biomedically relevant protein and organic ligands. Curr Med Chem. 2013;20(12):1595-608.

Folk JE, Piez KA, Carroll WR, Gladner JA. Carboxypeptidase B. Purification and characterization of the porcine enzyme. J Biol Chem. 1960;235:2272-7

García-Sáez I, Reverter D, Vendrell J, Avilés FX, Coll M. The three-dimensional structure of human procarboxypeptidase A2. Deciphering the basis of the inhibition, activation and intrinsic activity of the zymogen. EMBO J. 1997;16(23):6906-13.

Gomez-Ortiz M, Gomis-Rüth F, Huber R, Avilés F. Inhibition of carboxypeptidase A by excess zinc: analysis of the structural determinants by X-ray crystallography. FEBS Letters. 1997;400(3):336-340.

Gomis-Rüth F. Structure and Mechanism of Metallocarboxypeptidases. Crit Rev Bioch Mol Biol. 2008;43(5):319-45.

Pallarès I, Fernández D, Comellas-Bigler M, Fernández-Recio J, Ventura S, Avilés F et al. Direct interaction between a human digestive protease and the mucoadhesive poly(acrylic acid). Acta Crystallogr Section D Biol Crystallogr. 2008;64(7):784-791.

Bibliography

Rees DC, Lipscomb WN. Binding of ligands to the active site of carboxypeptidase A.Proc Natl Acad Sci U S A. 1981;78(9):5455-9.

Szmola R, Bence M, Carpentieri A, Szabo A, Costello C, Samuelson J et al. Chymotrypsin C Is a Co-activator of Human Pancreatic Procarboxypeptidases A1 and A2. J Biol Chem. 2011;286(3):1819-1827.

Valdez CE, Morgenstern A, Eberhart ME, Alexandrova AN. Predictive methods for computational metalloenzyme redesign - a test case with carboxypeptidase A. Phys Chem Chem Phys. 2016;18(46):31744-31756.

Vendrell J, Querol E, Avile FX. Metallocarboxypeptidases and their protein inhibitors: Structure, function and biomedical properties. Biochim Biophys Acta. 2000;1477(1-2):284-98.

Wu S, Zhang C, Xu D, Guo H. Catalysis of Carboxypeptidase A: Promoted-Water versus Nucleophilic Pathways. J Phys Chem B. 2010;114(28):9259-67.

Gomis-Rüth F. Structure and Mechanism of Metallocarboxypeptidases. Crit Rev Biochem Mol Biomol. 2008 Sep-Oct;43(5):319-45

Fernández D, Boix E, Pallarès I, Avilés FX, Vendrell J. Structural and Functional Analysis of the Complex between Citrate and the Zinc Peptidase Carboxypeptidase A. Enzyme Res. 2011;2011:128676

Lee Kj, Kim DH. Design of mechanism-based carboxypeptidase A inactivators on the basis of the X-ray crystal structure and catalytic reaction pathway. Bioorg Med Chem. 1998 Sep;6(9):1613-22

Fernández D, Pallarès I, Vendrell J, Avilés FX. Progress in metallocarboxypeptidases and their small molecular weight inhibitors. Biochimie. 2010 Nov;92(11):1484-500.

Bibliography

Fernández D, Testero S, Vendrell J, Avilés FX, Mobashery S. The X-ray structure of carboxypeptidase A inhibited by a thiirane mechanism-based inhibitor. Chem Biol Drug Des. 2010 Jan;75(1):29-34

Park JD, Kim DH, Kim SJ, Woo JR, Ryu SE. Sulfamide-based inhibitors for carboxypeptidase A. Novel type transition state analogue inhibitors for zinc proteases. J Med Chem. 2002 Nov 21;45(24):55295-302

Books

Berg J, Tymoczko J, Stryer L. Biochemistry. 1st ed. New York: W.H. Freeman and Co.; 2002.

Boyer P. Hydrolysis. 1st ed. New York: Academic Press; 1971.

Walker J. The Protein Protocols Handbook. 1st ed. Totowa, NJ: Humana Press; 2009.

Web sites

Carboxypeptidases A - MeSH - NCBI [Internet]. Ncbi.nlm.nih.gov. 2017 [cited 23 February 2017]. Available from: https://www.ncbi.nlm.nih.gov/mesh/68043422

EMBL-EBI I. Peptidase M14, carboxypeptidase A (IPR000834) < InterPro < EMBL-EBI [Internet]. Ebi.ac.uk. 2017 [cited 23 February 2017]. Available from: http://www.ebi.ac.uk/interpro/entry/IPR000834

MEROPS - the Peptidase Database [Internet]. Merops.sanger.ac.uk. 2017 [cited 23 February 2017]. Available from: http://merops.sanger.ac.uk/cgi-bin/famsum?family=m14

Bibliography

MeSH Browser [Internet]. Meshb.nlm.nih.gov. 2017 [cited 23 February 2017]. Available from: https://meshb.nlm.nih.gov/#/record/ui?name=Carboxypeptidases%20A

Metalloproteases [Internet]. Chemistry LibreTexts. 2017 [cited 23 February 2017]. Available from: https://chem.libretexts.org/Core/Biological_Chemistry/Catalysts/Metalloproteases

NCBI CDD Conserved Protein Domain M14_CPA [Internet]. Ncbi.nlm.nih.gov. 2017 [cited 23 February 2017]. Available from: https://www.ncbi.nlm.nih.gov/Structure/cdd/cddsrv.cgi?uid=cd03870

SCOP: Family: Pancreatic carboxypeptidases [Internet]. Scop.mrc-lmb.cam.ac.uk. 2017 [cited 23 February 2017]. Available from: http://scop.mrc-lmb.cam.ac.uk/scop/data/scop.b.d.hi.f.b.html

Multiple choice questions

Which characteristic is attributed to metallo-peptidases?

a. They have a metal element at the active site
b. They are endopeptidases
c. Both a and b are correct
d. They are exopeptidases
e. Both a and d are correct.

What is the difference between a pancreatic and a regulatory carboxypeptidase?

a. Regulatory carboxypeptidases are the digestive ones
b. Pancreatic carboxypeptidases are endopeptidases
c. Regulatory carboxypeptidases have zinc at their active site and pancreatic don’t
d. Pancreatic carboxypeptidases cleave proteins from diet
e. None of them is correct

Why are residues 69, 72 and 196 highly conserved?

a. Because they are acid residues
b. Because they are basic residues
c. Because they coordinate Zinc atom
d. Because they have long side chains
e. Because they are sulfur containing residues

How is Glu270 able to cleave the peptide bond?

a. Because it is able to act as a general acid allowing nucleophilic attack on the scissile amide carbon of the substrate
b. Because of its phosphate group
c. Because it binds to the substrate
d. Because it is a basic amino acid
e. Because it is an apolar amino acid

Multiple choice questions

About carboxypeptidase classification:

a. There is only one way to classify them, and it is very strict
b. There are different classifications, according to the criteria used
c. Both a and b are correct
d. A carboxypeptidase is always an endopeptidase
e. All of them are correct

When is a carboxypeptidase enzyme active?

a. When it is binded to the propeptide which blocks the active site
b. When the propeptide is a globular domain
c. Both a and b are correct
d. When it is not binded to the propetide which blocks the active site
e. None of them is correct

In which species carboxypeptidases are found?

a. Only in hummans
b. Only in mammals
c. Only in vertebrates
d. Only in eukaryotes
e. In eukaryotes and prokaryotes

How is Tyr 248 able to establish an hydrogen bond with the substrate?

a. Because it is an amino acid from the substrate
b. Because only aromatic amino acids perform hydrogen bonds
c. Both a and b are correct
d. Because it suffers a conformational change when the binding to the substrate occurs
e. All of them are correct

Multiple choice questions

About carboxypeptidases specificity:

1. Carboxypeptidase B has preference for basic residues.
2. Carboxypeptidase A1 cleaves aliphatic residues.
3. Carboxypeptidase A2 selectively acts on the bulkier aromatic residues.
4. All carboxypeptidases can cleave all the residues, there is no real specificity in these enzymes.

a. 1, 2, 3
b. 1, 3
c. 2, 4
d. 4
e. 1, 2 ,3 ,4

About sequence and structural similarity in carboxypeptidases:

1. Pancreatic carboxypeptidases are more similar among them that when compared to the regulatory ones
2. Similarities in sequence between digestive and non digestive carboxypeptidases are only 15-20%
3. Similarities in secondary structures between digestive and non digestive carboxypeptidases are higher than sequence similarities between them
4. Sequence is more similar between the two carboxypeptidase subfamilies than structure

a. 1, 2, 3
b. 1, 3
c. 2, 4
d. 4
e. 1, 2 ,3 ,4

Carboxypeptidases

Carlota Bellot Herrero

Marcel Lucas Sánchez

Irene Ortega González

Clàudia Prat Gibert