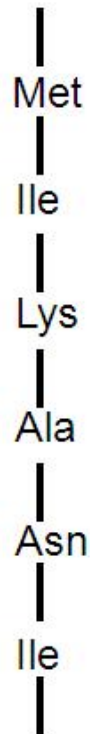


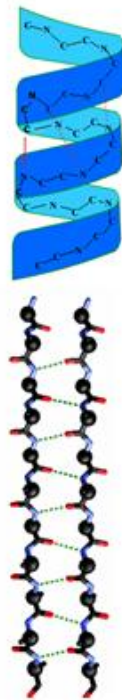
# Protein Structure

Primary  
structure



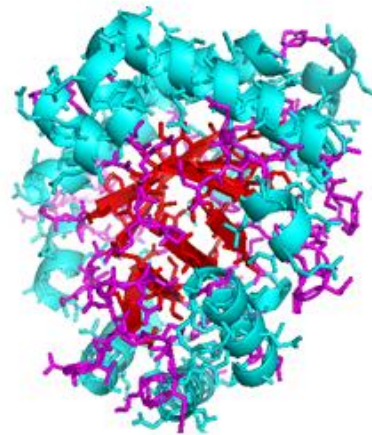
Amino acid sequence

Secondary  
structure



$\alpha$ -helix,  $\beta$ -strand

Tertiary  
structure



3D structure with  
atomic details

Quaternary  
structure



Assembled  
subunits

# Primary structure: human hemoglobin

>sp|P68871|HBB\_HUMAN Hemoglobin subunit beta OS=Homo sapiens

VHLTPEEKSAVTALWGKVNVDENVGGEALGRLLVYPWTQRFFESFGDLSTPD  
AVMGNPKVKAHGKKVLGAFSDGLAHLNLIKGTFTLSELHCDKLHVDPENFR  
LLGNVLVCVLAHHFG KEFTPPVQAAYQKVVAGVANALAHKYH

**Primary structure describes the linear sequence of amino acid residues in a protein.**

**It includes all covalent bonds between amino acids.**

**The relative arrangement of the linked amino acids is not specified.**

# Databases for protein sequences

EXProt

Munich Information Center for Protein Sequences (MIPS)

NCBI Protein database

**PIR - Protein Information Resource (Georgetown University)**

PIR-NREF

PRF

**SWISS-PROT (Swiss Institute of Bioinformatics)**

**TrEMBL**

**UniProt - The Universal Protein Knowledgebase**

# Protein Information Resource

**PRO:** Protein family classification

**iProClass:** integrated protein knowledgebase

**iProLink:** literature, information and knowledge

The screenshot shows the PIR website with a blue header. The PIR logo is on the left, followed by 'A UniProt CONORTIUM MEMBER' and 'Protein Information Resource'. On the right, there are three vertical text elements: 'TLALPM', 'LIGCLF', and 'VTGFSK'. Below the header is a navigation bar with links: 'About PIR', 'Databases', 'Search/Analysis', 'Download', and 'Support'. The main content area has a title 'INTEGRATED PROTEIN INFORMATICS RESOURCE FOR GENOMIC, PROTEOMIC AND SYSTEMS BIOLOGY RESEARCH'. Below this is a UniProt logo and a description: 'The Universal Protein Resource (UniProt) provides the scientific community with a single, centralized, authoritative resource for protein sequences and functional information.' It also lists 'UniProtKB | UniRef | UniParc' and 'Current release: 2011\_07'. The main content is divided into three columns. The first column is 'PRO Protein Ontology' with a graphic of yellow spheres connected by red lines. It lists: 'Representation of protein objects with descriptions and relationships', 'Browse PRO', and 'Annotate with RACE-PRO'. It also has a link for '\*Sample PRO report\*'. The second column is 'iProClass Integrated Protein Knowledgebase' with a graphic of blue cylinders. It lists: 'Value-added reports for UniProtKB and unique UniParc proteins' and 'Functional analysis and protein ID mapping'. It also has a link for '\*Sample protein report\*'. The third column is 'iProLINK Literature Information & Knowledge' with a graphic of a database cylinder and a molecular structure. It lists: 'Source for text mining and ontology development', 'RLIMS-P text mining tool, BioThesaurus', and 'Bibliography mapping'. It also has a link for '\*Sample Biblio. report\*'. Each column has a decorative border at the bottom.

PIR A UniProt CONORTIUM MEMBER Protein Information Resource

TLALPM  
LIGCLF  
VTGFSK

About PIR Databases Search/Analysis Download Support

INTEGRATED PROTEIN INFORMATICS RESOURCE FOR GENOMIC, PROTEOMIC AND SYSTEMS BIOLOGY RESEARCH

UniProt The Universal Protein Resource (UniProt) provides the scientific community with a single, centralized, authoritative resource for protein sequences and functional information.  
UniProtKB | UniRef | UniParc Current release: 2011\_07

**PRO** Protein Ontology

- Representation of protein objects with descriptions and relationships
- Browse PRO
- Annotate with RACE-PRO

\*Sample PRO report\*

**iProClass** Integrated Protein Knowledgebase

- Value-added reports for UniProtKB and unique UniParc proteins
- Functional analysis and protein ID mapping

\*Sample protein report\*

**iProLINK** Literature Information & Knowledge

- Source for text mining and ontology development
- RLIMS-P text mining tool, BioThesaurus
- Bibliography mapping

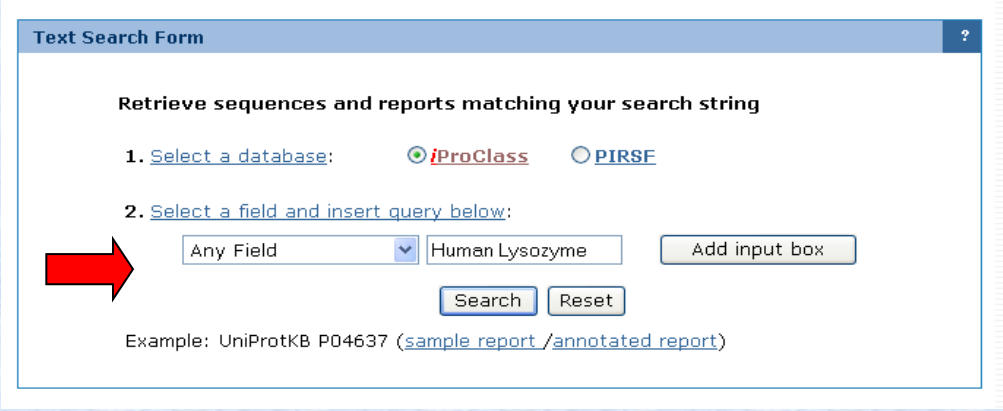
\*Sample Biblio. report\*

<http://pir.georgetown.edu/>



# Search with iProClass

The iProClass database provides value-added information reports on protein sequences, structures, families, functions, interactions, expressions and modifications.



The screenshot shows the 'Text Search Form' interface. It has a title bar with a question mark icon. The main content area is titled 'Retrieve sequences and reports matching your search string'. Below this, there are two numbered steps. Step 1, 'Select a database:', has two radio buttons: 'iProClass' (selected) and 'PIRSF'. Step 2, 'Select a field and insert query below:', features a dropdown menu with 'Any Field' selected, a text input box containing 'Human Lysozyme', and an 'Add input box' button. Below these are 'Search' and 'Reset' buttons. At the bottom, an example is provided: 'Example: UniProtKB P04637 (sample report / annotated report)'.

Text Search Form ?

Retrieve sequences and reports matching your search string

1. Select a database: ☒ iProClass ☐ PIRSF

2. Select a field and insert query below:

Any Field Human Lysozyme Add input box

Search Reset

Example: UniProtKB P04637 ([sample report](#) / [annotated report](#))

1 selected (show)		GO Slim / KEGG Pathway...		BLAST FASTA Pattern Match Pairwise Alignment Multiple Alignment Domain Display		
<input type="checkbox"/> Protein AC/ID	Protein Name	Length	Organism Name	PIRSF ID	Related Seq. +	Matched Fields
<input type="checkbox"/> P79239/LYSC_PONPY <small>/ProClass UniProtKB/Swiss-Prot BioThesaurus</small>	Lysozyme C precursor	148	<a href="#">Pongo pygmaeus (Bornean orangutan)</a>	<a href="#">PIRSF001064</a>	<a href="#">300</a>	Paper Title=>human lysozyme
<input type="checkbox"/> P61628/LYSC_PANTR <small>/ProClass UniProtKB/Swiss-Prot BioThesaurus</small>	Lysozyme C precursor	148	<a href="#">Pan troglodytes (Chimpanzee)</a>	<a href="#">PIRSF001064</a>	<a href="#">300</a>	Paper Title=>human lysozyme
<input checked="" type="checkbox"/> P61627/LYSC_PANPA <small>/ProClass UniProtKB/Swiss-Prot BioThesaurus</small>	Lysozyme C precursor	148	<a href="#">Pan paniscus (Pygmy chimpanzee) (Bonobo)</a>	<a href="#">PIRSF001064</a>	<a href="#">300</a>	Paper Title=>human lysozyme
<input type="checkbox"/> P61626/LYSC_HUMAN <small>/ProClass UniProtKB/Swiss-Prot BioThesaurus</small>	Lysozyme C precursor	148	<a href="#">Homo sapiens (Human)</a>	<a href="#">PIRSF001064</a>	<a href="#">300</a>	Paper Title=>human lysozyme; Paper Title=>human
<input type="checkbox"/> P79179/LYSC_GORGO <small>/ProClass UniProtKB/Swiss-Prot BioThesaurus</small>	Lysozyme C precursor	148	<a href="#">Gorilla gorilla gorilla (Lowland gorilla)</a>	<a href="#">PIRSF001064</a>	<a href="#">300</a>	Paper Title=>human lysozyme
<input type="checkbox"/> P02788/TRFL_HUMAN <small>/ProClass UniProtKB/Swiss-Prot BioThesaurus</small>	Lactotransferrin precursor	710	<a href="#">Homo sapiens (Human)</a>	<a href="#">PIRSF002549;</a> <a href="#">PIRSF500683</a>	<a href="#">300</a>	Paper Title=>human lysozyme
<input type="checkbox"/> Q6PCD2/Q6PCD2_HUMAN <small>/ProClass UniProtKB/TrEMBL BioThesaurus</small>	GABRE protein	365	<a href="#">Homo sapiens (Human)</a>		<a href="#">300</a>	Paper Title=>human lysozyme
<input type="checkbox"/> B2R4C5/B2R4C5_HUMAN <small>/ProClass UniProtKB/TrEMBL BioThesaurus</small>	Lysozyme (Renal amyloidosis), isoform CRA_a	148	<a href="#">Homo sapiens (Human)</a>		<a href="#">300</a>	Paper Title=>human lysozyme; Paper Title=>human
<input type="checkbox"/> Q876Z9/Q876Z9_ASPOR <small>/ProClass UniProtKB/TrEMBL BioThesaurus</small>	Predicted protein	600	<a href="#">Aspergillus oryzae</a>	<a href="#">PIRSF037788;</a> <a href="#">PIRSF500676</a>	<a href="#">300</a>	Paper Title=>human lysozyme
<input type="checkbox"/> Q4R8K7/Q4R8K7_MACFA <small>/ProClass UniProtKB/TrEMBL BioThesaurus</small>	Testis cDNA clone: QtsA-12244, similar to human lysozyme homolog (LOC57151),	109	<a href="#">Macaca fascicularis (Crab eating macaque)</a> <a href="#">Cynomolgus monkey</a>		<a href="#">300</a>	Protein Name=>human lysozyme

Save Result As: [TABLE](#) | [FASTA](#)

1 selected (show)

GO Slim / KEGG Pathway...

[BLAST](#) [FASTA](#) [Pattern Match](#) [Pairwise Alignment](#) [Multiple Alignment](#) [Domain Display](#)

<input checked="" type="checkbox"/> Protein AC/ID	Protein Name	Length	Organism Name	PIRSF ID	Related Seq. +	Matched Fields
<input checked="" type="checkbox"/> P61626/LYSC_HUMAN <a href="#">ProClass</a> <a href="#">UniProtKB/Swiss-Prot</a> <a href="#">BioThesaurus</a>	Lysozyme C precursor	148	<a href="#">Homo sapiens (Human)</a>	<a href="#">PIRSF001064</a>	<a href="#">300</a>	UniProtKB AC=>P61626

**ProClass Summary Report for UniProtKB Entry: P61626**

[Related Sequences](#)

[BioThesaurus](#)

[ID Mapping](#)

**GENERAL INFORMATION**

Protein Name and ID	UniProtKB ID	UniProtKB Accession	Protein Name
	<a href="#">LYSC_HUMAN</a>	<a href="#">P61626</a> ; P00695; Q13170; Q9UCF8	Lysozyme C precursor
PIR-PSD: <a href="#">LZHU</a> RefSeq: <a href="#">NP_000230.1</a> GenPept: <a href="#">AAA59535.1</a> ; <a href="#">AAC63078.1</a> ; <a href="#">FAW97222.1</a> ; <a href="#">AAA59536.1</a> ; <a href="#">CAA32175.1</a> ; <a href="#">AAH04147.1</a> ; <a href="#">FAW97221.1</a> ; <a href="#">ACO37637.1</a> ; <a href="#">AAA36188.1</a> IPI: <a href="#">IPI00019038</a>			
Taxonomy	<i>Source Organism:</i> Homo sapiens (Human) <i>Taxon Group:</i> Euk/mammal <i>NCBI Taxon:</i> <a href="#">9606</a> <i>Lineage:</i> Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo.		
Gene Name	LYZ; LZM		
Keywords	3d-structure; amyloid; amyloidosis; antimicrobial; bacteriolytic enzyme; direct protein sequencing; disease mutation; disulfide bond; glycosidase; hydrolase; polymorphism; polysaccharide degradation; signal		
Function	Lysozymes have primarily a bacteriolytic function; those in tissues and body fluids are associated with the monocyte- macrophage system and enhance the activity of immunoagents.		
Subunit	Monomer.		

**CROSS-REFERENCES**

Bibliography	<a href="#">►View Bibliography Information</a> <a href="#">►Submit Bibliography</a> <i>Annotated references:</i> PMID: <a href="#">8105095</a> ; <a href="#">10350481</a> ; <a href="#">10469827</a> ; <a href="#">10561612</a> ; <a href="#">11887182</a> ; <a href="#">11927576</a> ; <a href="#">11986950</a> [PDB/GeneRIF] <a href="#">More</a>
	<i>Other references:</i> PMID: <a href="#">11849445</a> ; <a href="#">12675840</a> ; <a href="#">15745733</a> ; <a href="#">8765309</a> ; <a href="#">9659355</a> ; <a href="#">9745729</a> ; <a href="#">18391951</a> ; <a href="#">9359845</a> ; <a href="#">8566845</a> ; <a href="#">17353931</a> ; <a href="#">9883972</a> ; <a href="#">366724</a> ; <a href="#">10534505</a> ; <a href="#">12477932</a> ; <a href="#">10558865</a> ; <a href="#">18591461</a>
DNA Sequence	GenBank/EMBL/DDBJ: <a href="#">M21119</a> ; <a href="#">J03801</a> ; <a href="#">M19045</a> ; <a href="#">X14008</a> ; <a href="#">U25677</a> ; <a href="#">BC004147</a>

Structure	<div>1B5U: <a href="#">SCOP</a> <a href="#">CATH</a> <a href="#">FSSP</a> <a href="#">MMDb</a> <a href="#">PDBsum</a></div> <div>1B5V: <a href="#">SCOP</a> <a href="#">CATH</a> <a href="#">FSSP</a> <a href="#">MMDb</a> <a href="#">PDBsum</a></div> <div>1B5W: <a href="#">SCOP</a> <a href="#">CATH</a> <a href="#">FSSP</a> <a href="#">MMDb</a> <a href="#">PDBsum</a></div> <div>1B5X: <a href="#">SCOP</a> <a href="#">CATH</a> <a href="#">FSSP</a> <a href="#">MMDb</a> <a href="#">PDBsum</a></div> <div>1B5Y: <a href="#">SCOP</a> <a href="#">CATH</a> <a href="#">FSSP</a> <a href="#">MMDb</a> <a href="#">PDBsum</a></div> <div>1B5Z: <a href="#">SCOP</a> <a href="#">CATH</a> <a href="#">FSSP</a> <a href="#">MMDb</a> <a href="#">PDBsum</a></div> <div>1B7L: <a href="#">SCOP</a> <a href="#">CATH</a> <a href="#">FSSP</a> <a href="#">MMDb</a> <a href="#">PDBsum</a></div> <div>1B7M: <a href="#">SCOP</a> <a href="#">CATH</a> <a href="#">FSSP</a> <a href="#">MMDb</a> <a href="#">PDBsum</a></div> <div>1B7N: <a href="#">SCOP</a> <a href="#">CATH</a> <a href="#">FSSP</a> <a href="#">MMDb</a> <a href="#">PDBsum</a></div> <div>1B7O: <a href="#">SCOP</a> <a href="#">CATH</a> <a href="#">FSSP</a> <a href="#">MMDb</a> <a href="#">PDBsum</a></div> <div>1B7P: <a href="#">SCOP</a> <a href="#">CATH</a> <a href="#">FSSP</a> <a href="#">MMDb</a> <a href="#">PDBsum</a></div> <div>1B7Q: <a href="#">SCOP</a> <a href="#">CATH</a> <a href="#">FSSP</a> <a href="#">MMDb</a> <a href="#">PDBsum</a></div> <div>1B7R: <a href="#">SCOP</a> <a href="#">CATH</a> <a href="#">FSSP</a> <a href="#">MMDb</a> <a href="#">PDBsum</a></div> <div>1B7S: <a href="#">SCOP</a> <a href="#">CATH</a> <a href="#">FSSP</a> <a href="#">MMDb</a> <a href="#">PDBsum</a></div> <div>1B83: <a href="#">SCOP</a> <a href="#">CATH</a> <a href="#">FSSP</a> <a href="#">MMDb</a> <a href="#">PDBsum</a></div> <div>1B84: <a href="#">SCOP</a> <a href="#">CATH</a> <a href="#">FSSP</a> <a href="#">MMDb</a> <a href="#">PDBsum</a></div> <div>More</div>
PIR Feature & Post Translational Modifications	<div>FEAT1; active site: Glu, Asp (53,71) [predicted]</div> <div>FEAT2; binding site: substrate (Asp) (120) [predicted]</div> <div>FEAT3; disulfide bonds: (24-146,48-134,83-99,95-113) [experimental]</div> <div>FEAT4; domain: signal sequence (1-18) [predicted]</div> <div>FEAT5; product: lysozyme (19-148) [experimental]</div> <div>Phosphosite: <a href="#">P61626</a></div>

FAMILY CLASSIFICATION	
UniRef	<a href="#">UniRef100_P61626</a> ; <a href="#">UniRef90_P61626</a> ; <a href="#">UniRef50_P61626</a>
PIRSF	<a href="#">PIRSF001064</a> lysozyme c
Pfam Domain	Pfam: <a href="#">PF00062</a> : C-type lysozyme/alpha-lactalbumin family (19-146)
Prosite Motif	Prosite: <a href="#">PS00128</a> : <a href="#">PDQC00119</a> : Alpha-lactalbumin / lysozyme C signature. Prosite: <a href="#">PS51348</a> : <a href="#">PDQC00119</a> : Alpha-lactalbumin / lysozyme C family profile.
InterPro	InterPro: <a href="#">LYSC_HUMAN</a> <a href="#">IPR001916</a> : Glycoside hydrolase, family 22 <a href="#">IPR000974</a> : Glycoside hydrolase, family 22, lysozyme
SCOP Fold	►Class: <a href="#">Alpha and beta proteins (a+b)</a> ; Fold: <a href="#">Lysozyme-like</a> ; Superfamily: <a href="#">Lysozyme-like</a> ; Family: <a href="#">C-type lysozyme</a> [133L:A; 134L:A; 1B5U:A; 1B5V:A; 1B5W:A; 1B5X:A; 1B5Y:A; 1B5Z:A; 1B5Z:B; 1B7L:A; 1B7M:A; 1B7N:A; 1B7O:A; 1B7P:A; 1B7Q:A; 1B7R:A; 1B7S:A; 1B83:A; 1BB3:B; 1BB4:A; 1BB4:B; 1BB5:A; 1BB5:B; 1C43:A; 1C45:A; 1C46:A; 1C7P:A; 1CJ6:A; 1CJ7:A; 1CJ8:A; 1CJ9:A; 1CKC:A; 1CKD:A; 1CKF:A; 1CKG:A; 1CKH:A; 1D6P:A; 1D6Q:A; 1DI3:A; 1DI4:A; 1DI5:A; 1EQ4:A; 1EQ5:A; 1EQE:A; 1GAY:A; 1GAZ:A; 1GB0:A; 1GB2:A; 1M9E]
Other Classification	BLOCKS: <a href="#">IPB000974</a> Lysozyme signature PRINTS: <a href="#">PR00137</a> LYSOZYME PRINTS: <a href="#">PR00135</a> LYZLACT SMART: <a href="#">SM00263</a> LYZ1

FEATURE & SEQUENCE DISPLAY	
	<div><div><div>Length = 148</div><div>P61626</div><div>PF00062</div><div>1</div><div>61</div><div>121</div></div><div><div>Click on a bar to show its sequence; to copy and paste it, press ctrl-c then ctrl-v.</div><div><div>1</div><div>148</div></div><div><div>1</div><div>2</div><div>3</div><div>4</div><div>5</div><div>6</div><div>7</div><div>8</div><div>9</div><div>10</div><div>11</div><div>12</div><div>13</div><div>14</div><div>15</div><div>16</div><div>17</div><div>18</div><div>19</div><div>20</div><div>21</div><div>22</div><div>23</div><div>24</div><div>25</div><div>26</div><div>27</div><div>28</div><div>29</div><div>30</div><div>31</div><div>32</div><div>33</div><div>34</div><div>35</div><div>36</div><div>37</div><div>38</div><div>39</div><div>40</div><div>41</div><div>42</div><div>43</div><div>44</div><div>45</div><div>46</div><div>47</div><div>48</div><div>49</div><div>50</div><div>51</div><div>52</div><div>53</div><div>54</div><div>55</div><div>56</div><div>57</div><div>58</div><div>59</div><div>60</div><div>61</div><div>62</div><div>63</div><div>64</div><div>65</div><div>66</div><div>67</div><div>68</div><div>69</div><div>70</div><div>71</div><div>72</div><div>73</div><div>74</div><div>75</div><div>76</div><div>77</div><div>78</div><div>79</div><div>80</div><div>81</div><div>82</div><div>83</div><div>84</div><div>85</div><div>86</div><div>87</div><div>88</div><div>89</div><div>90</div><div>91</div><div>92</div><div>93</div><div>94</div><div>95</div><div>96</div><div>97</div><div>98</div><div>99</div><div>100</div><div>101</div><div>102</div><div>103</div><div>104</div><div>105</div><div>106</div><div>107</div><div>108</div><div>109</div><div>110</div><div>111</div><div>112</div><div>113</div><div>114</div><div>115</div><div>116</div><div>117</div><div>118</div><div>119</div><div>120</div><div>121</div><div>122</div><div>123</div><div>124</div><div>125</div><div>126</div><div>127</div><div>128</div><div>129</div><div>130</div><div>131</div><div>132</div><div>133</div><div>134</div><div>135</div><div>136</div><div>137</div><div>138</div><div>139</div><div>140</div><div>141</div><div>142</div><div>143</div><div>144</div><div>145</div><div>146</div><div>147</div><div>148</div></div><div><div>MKALIVLGLVLLSVTVQGGKVFERCELA</div><div>RTTLKRLGMDGGYRGISLANWMCLAKWESGYNTRA</div><div>TNYNAGDRSTDYGI</div><div>FQINSRYWCNDGKTPGAVNACHLSCSALLQDNIADAVACAKRVVRD</div><div>PQGIRAWVAWRNRCQNRDVRQYVQGGCV</div></div></div></div>



# Swiss-prot/Uniprot

Annotated protein sequence database established in 1986 and maintained collaboratively, since 1987, by the Department of Medical Biochemistry of the **University of Geneva** and the **EMBL** Data Library.

It is a curated protein sequence database which strives to provide a **high level of annotation** (such as the description of the function of a protein, its domain structure, post-translational modifications and variants), a **minimal level of redundancy** and a **high level of integration** with other databases.

TrEMBL is a **computer annotated** supplement of SWISS-PROT that contains all the translations of EMBL nucleotide sequence entries not yet integrated in SWISS-PROT.

Currently, SWISS-PROT and TrEMBL have 0.55 and 73.7 million sequences, respectively.

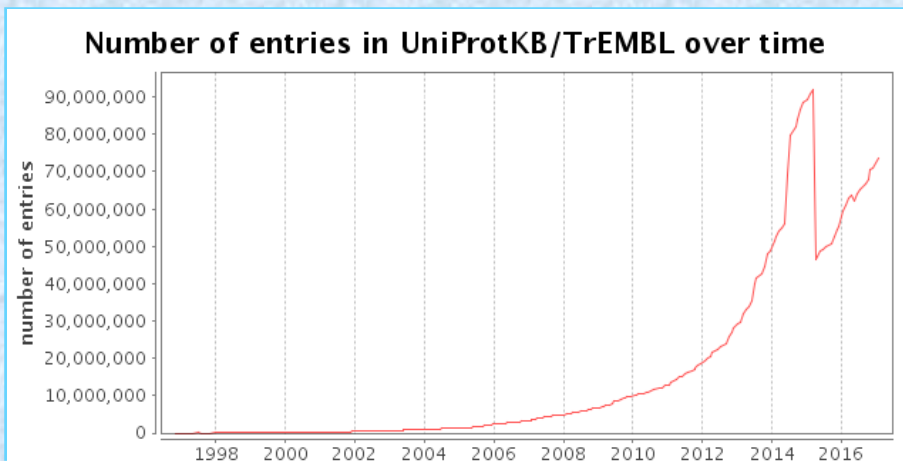
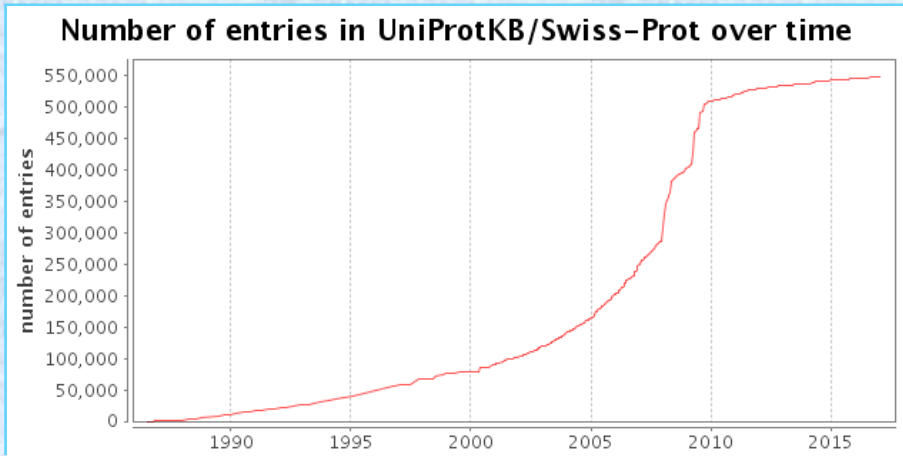
Total: 74.4 million

<http://www.ebi.ac.uk/swissprot/>

<http://www.uniprot.org/uniprot/>

# Uniprot: statistics

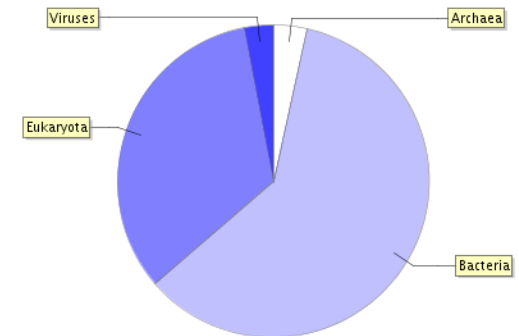
Number of entries: ~74.4 million



## Taxonomic Origin

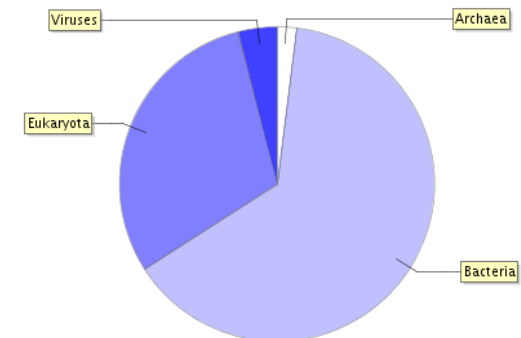
Swiss-Prot entries per taxonomic group

Swiss-Prot entries per taxonomic group



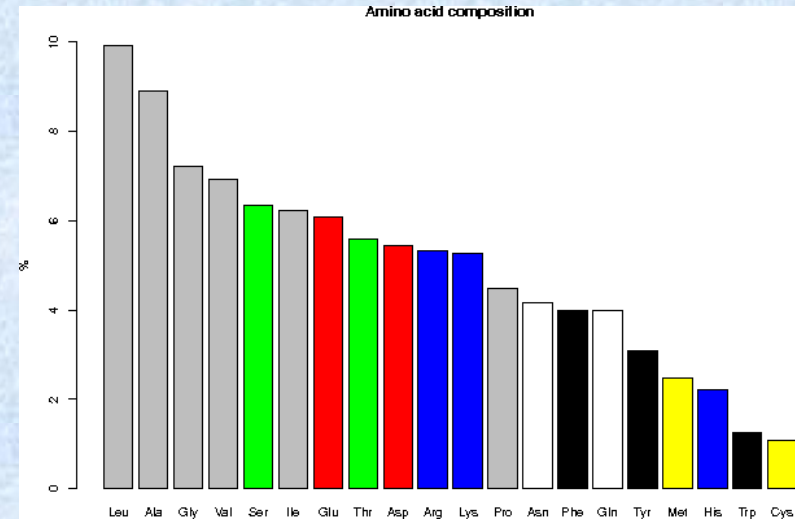
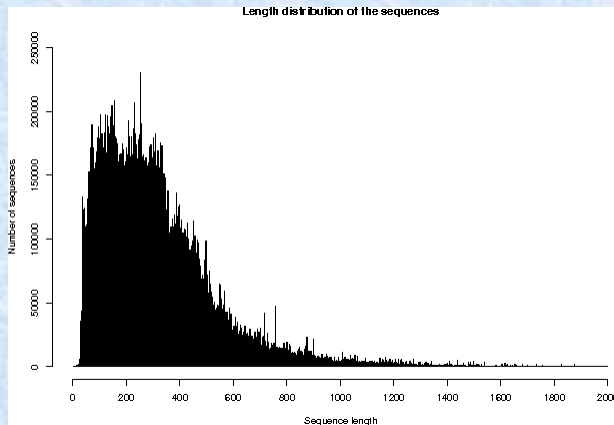
## Taxonomic Origin

TrEMBL entries per taxonomic group

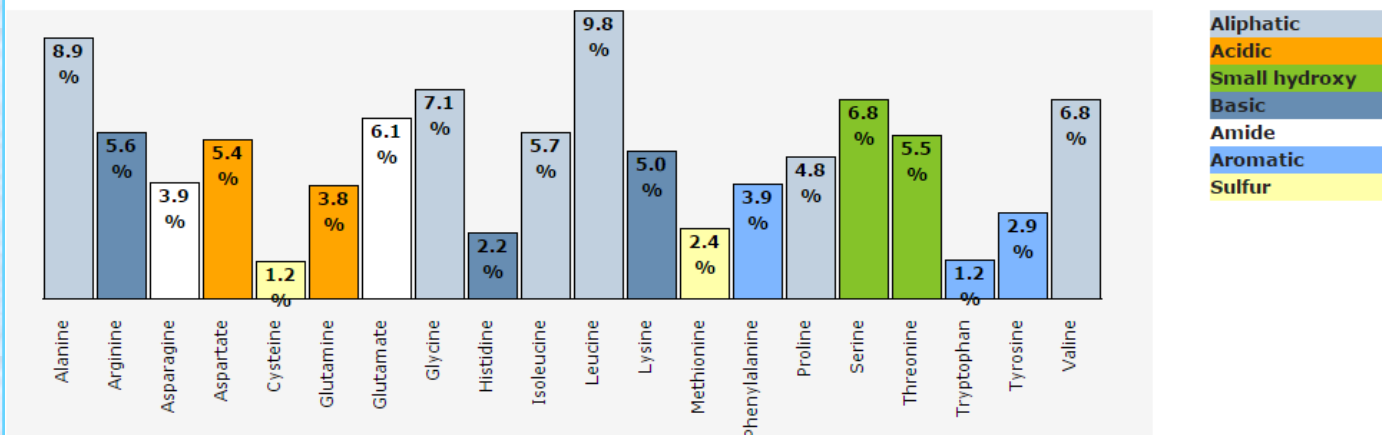


# Uniprot: statistics

Average sequence length:  
315 amino acids



## Amino acid distribution statistics



# Uniprot: contents

**Name and origin of the protein**

**protein attributes**

**general information**

**Ontologies**

**sequence annotation**

**amino acid sequence**

**bibliographic references**

**cross-references with sequence, structure and interaction databases**

**entry information.**

# Uniprot: search results

UniProt > UniProtKB

Search Blast Align Retrieve ID Mapping \*

Search in **Query**

Protein Knowledgebase (UniProtKB) ☒ hemoglobin B chain

Search Advanced Search » Clear

1 - 25 of 1,243 results for **hemoglobin** AND **B** AND **chain** in UniProtKB sorted by **score** descending

Browse by [taxonomy](#), [keyword](#), [gene ontology](#), [enzyme class](#) or [pathway](#) | Reduce sequence redundancy to 100%, 90% or 50% |

Download

Page 1

of 50 | Next »

## Results [Customize](#)

> Show only [reviewed \(825\)](#) ★ (UniProtKB/Swiss-Prot) or [unreviewed \(418\)](#) ★ (UniProtKB/TrEMBL) entries

> Quote terms: "hemoglobin b"

> Restrict term "hemoglobin" to [protein family \(24\)](#), [gene name \(3\)](#), [gene ontology \(883\)](#), [keyword \(9\)](#), [protein name \(970\)](#), [web resource \(5\)](#)

> Restrict term "b" to [author \(608\)](#), [domain \(4\)](#), [gene name \(15\)](#), [gene ontology \(5\)](#), [keyword \(4\)](#), [protein name \(58\)](#), [organism \(1\)](#), [source \(7\)](#), [strain \(5\)](#), [taxonomy \(6\)](#), [tissue \(2\)](#)

> Restrict term "chain" to [author \(48\)](#), [gene name \(2\)](#), [gene ontology \(7\)](#), [keyword \(7\)](#), [protein name \(740\)](#), [annotation topic \(943\)](#)

> Restrict term "chain" to [pathway](#)

Accession	Entry name	Status	Protein names	Gene names	Organism	Length
<input checked="" type="checkbox"/> <a href="#">P68871</a>	HBB_HUMAN	★	Hemoglobin subunit beta	HBB	Homo sapiens (Human)	147
<input type="checkbox"/> <a href="#">P69905</a>	HBA_HUMAN	★	Hemoglobin subunit alpha	HBA1 HBA2	Homo sapiens (Human)	142
<input type="checkbox"/> <a href="#">P69892</a>	HBG2_HUMAN	★	Hemoglobin subunit gamma-2	HBG2	Homo sapiens (Human)	147
<input type="checkbox"/> <a href="#">P69891</a>	HBG1_HUMAN	★	Hemoglobin subunit gamma-1	HBG1 PRO2979	Homo sapiens (Human)	147
<input type="checkbox"/> <a href="#">P0A592</a>	TRHBN_MYCTU	★	Group 1 truncated hemoglobin glbN	glbN Rv1542c MT1594 MTCY48.23	Mycobacterium tuberculosis	136
<input type="checkbox"/> <a href="#">P02042</a>	HBD_HUMAN	★	Hemoglobin subunit delta	HBD	Homo sapiens (Human)	147



# Uniprot: contents

## Names and origin

Protein names	<i>Recommended name:</i> <b>Hemoglobin subunit beta</b> <i>Alternative name(s):</i> Beta-globin Hemoglobin beta chain <u>Cleaved into the following chain:</u> 1. <a href="#">LVV-hemorphin-7</a>
Gene names	Name: <b>HBB</b>
Organism	<a href="#">Homo sapiens (Human)</a> [Complete proteome]
Taxonomic identifier	<a href="#">9606</a> [NCBI]
Taxonomic lineage	<a href="#">Eukaryota</a> > <a href="#">Metazoa</a> > <a href="#">Chordata</a> > <a href="#">Craniata</a> > <a href="#">Vertebrata</a> > <a href="#">Euteleostomi</a> > <a href="#">Mammalia</a> > <a href="#">Eutheria</a> > <a href="#">Euarchontoglires</a> > <a href="#">Primates</a> > <a href="#">Haplorrhini</a> > <a href="#">Catarrhini</a> > <a href="#">Hominidae</a> > <a href="#">Homo</a>

## Protein attributes

Sequence length	147 AA.
Sequence status	Complete.
Sequence processing	The displayed sequence is further processed into a mature form.
Protein existence	Evidence at protein level.

# Uniprot: contents

## General annotation (Comments)

Function	Involved in oxygen transport from the lung to the various peripheral tissues. <a href="#">Ref.35</a> LVV-hemorphin-7 potentiates the activity of bradykinin, causing a decrease in blood pressure. <a href="#">Ref.35</a>
Subunit structure	Heterotetramer of two alpha chains and two beta chains in adult hemoglobin A (HbA).
Tissue specificity	Red blood cells.
Post-translational modification	Glucose reacts non-enzymatically with the N-terminus of the beta chain to form a stable ketoamine linkage. This takes place slowly and continuously throughout the 120-day life span of the red blood cell. The rate of glycation is increased in patients with diabetes mellitus. S-nitrosylated; a nitric oxide group is first bound to Fe <sup>2+</sup> and then transferred to Cys-94 to allow capture of O <sub>2</sub> . Acetylated on Lys-60, Lys-83 and Lys-145 upon aspirin exposure. <a href="#">Ref.34</a> reports the identification of HBB acetylated on Lys-145 in the cytosolic fraction of HeLa cells. This may have resulted from contamination of the sample.
Involvement in disease	Defects in HBB may be a cause of Heinz body anemias (HEIBAN) [MIM:140700]. This is a form of non-spherocytic hemolytic anemia of Dacie type 1. After splenectomy, which has little benefit, basophilic inclusions called Heinz bodies are demonstrable in the erythrocytes. Before splenectomy, diffuse or punctate basophilia may be evident. Most of these cases are probably instances of hemoglobinopathy. The hemoglobin demonstrates heat lability. Heinz bodies are observed also with the Ivemark syndrome (asplenia with cardiovascular anomalies) and with glutathione peroxidase deficiency. <a href="#">Ref.51</a> <a href="#">Ref.125</a> <a href="#">Ref.126</a> <a href="#">Ref.129</a>  Defects in HBB are the cause of beta-thalassemia (B-THAL) [MIM:604131]. A form of thalassemia. Thalassemias are common monogenic diseases occurring mostly in Mediterranean and Southeast Asian populations. The hallmark of beta-thalassemia is an imbalance in globin-chain production in the adult HbA molecule. Absence of beta chain causes beta(0)-thalassemia, while reduced amounts of detectable beta globin causes beta <sup>+</sup> -thalassemia. In the severe forms of beta-thalassemia, the excess alpha globin chains accumulate in the developing erythroid precursors in the marrow. Their deposition leads to a vast increase in erythroid apoptosis that in turn causes ineffective erythropoiesis and severe microcytic hypochromic anemia. Clinically, beta-thalassemia is divided into thalassemia major which is transfusion dependent, thalassemia intermedia (of intermediate severity), and thalassemia minor that is asymptomatic. <a href="#">Ref.52</a>  Defects in HBB are the cause of sickle cell anemia (SKCA) [MIM:603903]; also known as sickle cell disease. Sickle cell anemia is characterized by abnormally shaped red cells resulting in chronic anemia and periodic episodes of pain, serious infections and damage to vital organs. Normal red blood cells are round and flexible and flow easily through blood vessels, but in sickle cell anemia, the abnormal hemoglobin (called Hb S) causes red blood cells to become stiff. They are C-shaped and resembles a sickle. These stiffer red blood cells can lead to microvascular occlusion thus cutting off the blood supply to nearby tissues.  Defects in HBB are the cause of beta-thalassemia dominant inclusion body type (B-THALIB) [MIM:603902]. An autosomal dominant form of beta thalassemia characterized by moderate anemia, lifelong jaundice, cholelithiasis and splenomegaly, marked morphologic changes in the red cells, erythroid hyperplasia of the bone marrow with increased numbers of multinucleate red cell precursors, and the presence of large inclusion bodies in the normoblasts, both in the marrow and in the peripheral blood after splenectomy. <a href="#">Ref.52</a>
Miscellaneous	One molecule of 2,3-bisphosphoglycerate can bind to two beta chains per hemoglobin tetramer.

# Uniprot: contents

## Ontologies

### Keywords

Biological process	Oxygen transport Transport
Coding sequence diversity	Polymorphism
Disease	Congenital dyserythropoietic anemia Disease mutation Hereditary hemolytic anemia
Ligand	Heme Iron Metal-binding Pyruvate
Molecular function	Hypotensive agent Vasoactive
PTM	Acetylation Glycation Glycoprotein Phosphoprotein S-nitrosylation
Technical term	3D-structure Complete proteome Direct protein sequencing

## Gene Ontology (GO)

### Biological process

#### [blood coagulation](#)

Traceable author statement. Source: Reactome

#### [hydrogen peroxide catabolic process](#)

Inferred from direct assay. Source: BHF-UCL

#### [nitric oxide transport](#)

Non-traceable author statement. Source: UniProtKB

#### [positive regulation of cell death](#)

Inferred from direct assay. Source: BHF-UCL

#### [positive regulation of nitric oxide biosynthetic process](#)

Non-traceable author statement. Source: UniProtKB

#### [protein heterooligomerization](#)

Inferred from direct assay. Source: BHF-UCL

#### [regulation of blood pressure](#)

Inferred from electronic annotation. Source: UniProtKB-KW

#### [regulation of blood vessel size](#)

Inferred from electronic annotation. Source: UniProtKB-KW

### Cellular component

#### [haptoglobin-hemoglobin complex](#)

Inferred from direct assay. Source: BHF-UCL

#### [hemoglobin complex](#)

Non-traceable author statement [Ref.33](#) [Ref.71](#). Source: UniProtKB

### Molecular function

#### [heme binding](#)

Inferred from electronic annotation. Source: InterPro

#### [hemoglobin binding](#)

Inferred from direct assay. Source: UniProtKB

#### [oxygen binding](#)

Inferred from direct assay. Source: UniProtKB

#### [oxygen transporter activity](#)

Non-traceable author statement [Ref.71](#). Source: UniProtKB

# Uniprot: contents

Binary interactions									
With	Entry	#Exp.	IntAct	No	Natural variations				
HBA1	P69905	1	EBI-715554,EBI-714680		<input type="checkbox"/>	Natural variant	2	1	V → A in Raleigh; O(2) affinity down. [dbSNP:rs33949930]
					<input type="checkbox"/>	Natural variant	3	1	H → L in Graz. [dbSNP:rs35906307] <a href="#">Ref.76</a>
					<input type="checkbox"/>	Natural variant	3	1	H → Q in Okayama; O(2) affinity up. [dbSNP:rs713040]
					<input type="checkbox"/>	Natural variant	3	1	H → R in Deer Lodge; O(2) affinity up. [dbSNP:rs33983205]
					<input type="checkbox"/>	Natural variant	3	1	H → Y in Fukuoka. [dbSNP:rs35906307]
					<input type="checkbox"/>	Natural variant	6	1	P → R in Warwickshire. [dbSNP:rs34769005]
					<input type="checkbox"/>	Natural variant	7	1	E → A in G-Makassar.
					<input type="checkbox"/>	Natural variant	7	1	E → K in C. <a href="#">Ref.3</a> <a href="#">Ref.49</a>
					<input type="checkbox"/>	Natural variant	7	1	E → Q in Machida. [dbSNP:rs33930165]
					<input type="checkbox"/>	Natural variant	7	1	E → V in S; sickle cell anemia. [dbSNP:rs334] <a href="#">Ref.10</a> <a href="#">Ref.39</a>
					<input type="checkbox"/>	Natural variant	8	1	E → G in G-San Jose; mildly unstable. [dbSNP:rs34948328]
					<input type="checkbox"/>	Natural variant	8	1	E → K in G-Siriraj. [dbSNP:rs34948328]
					<input type="checkbox"/>	Natural variant	9	1	K → E in N-Timone. [dbSNP:rs33932981] <a href="#">Ref.101</a>
					<input type="checkbox"/>	Natural variant	9	1	K → Q in J-Luhe. [dbSNP:rs33926764]
					<input type="checkbox"/>	Natural variant	9	1	K → T in Rio Grande. <a href="#">Ref.114</a>
					<input type="checkbox"/>	Natural variant	10	1	S → C in Porto Alegre; O(2) affinity up. [dbSNP:rs33918131]
					<input type="checkbox"/>	Natural variant	11	1	A → D in Ankara. [dbSNP:rs33947457] <a href="#">Ref.55</a>
Sequence annotation (Features)									
	Feature key	Position(s)	Length	Description					
<b>Molecule processing</b>									
<input type="checkbox"/>	Initiator methionine	1	1	Removed					
<input checked="" type="checkbox"/>	Chain	2 – 147	146	Hemoglobin					
<input type="checkbox"/>	Peptide	33 – 42	10	LVV-he					
<b>Sites</b>									
<input type="checkbox"/>	Metal binding	64	1	Iron (h					
<input type="checkbox"/>	Metal binding	93	1	Iron (h					
<input type="checkbox"/>	Binding site	2	1	2,3-bis					
<input type="checkbox"/>	Binding site	3	1	2,3-bis					
<input type="checkbox"/>	Binding site	83	1	2,3-bis					
<input type="checkbox"/>	Binding site	144	1	2,3-bis					

## Cross-references

### Sequence databases

- ☒ EMBL
- ☐ GenBank
- ☐ DDBJ

ProteinModelPortal	<a href="#">P68871.</a>
SMR	<a href="#">P68871.</a> Positions <a href="#">2-147.</a>
ModBase	<a href="#">Search...</a>

### Protein-protein interaction databases

IntAct	<a href="#">P68871.</a> 24 interactions.
MINT	<a href="#">MINT-5000306.</a>
STRING	<a href="#">P68871.</a>

### PTM databases

PhosphoSite	<a href="#">P68871.</a>
-------------	-------------------------

### 2-D gel databases

SWISS-2DPAGE	<a href="#">P68871.</a>
PMMA-2DPAGE	<a href="#">P68871.</a>
REPRODUCTION-2DPAGE	<a href="#">IP100654755.</a> <a href="#">P68871.</a>
Siena-2DPAGE	<a href="#">P68871.</a>
UCD-2DPAGE	<a href="#">P02023.</a> <a href="#">P68871.</a>

### Proteomic databases

PeptideAtlas	<a href="#">P68871.</a>
PRIDE	<a href="#">P68871.</a>

### Secondary structure



[Details...](#)

### 3D structure databases

- ☒ PDBe
- ☐ RCSB PDB
- ☐ PDBj


Entry	Method	Resolution
<a href="#">1A00</a>	X-ray	2.50
<a href="#">1A01</a>	X-ray	2.50
<a href="#">1A0U</a>	X-ray	2.50
<a href="#">1A0Z</a>	X-ray	2.50
<a href="#">1A3N</a>	X-ray	2.50
<a href="#">1A3O</a>	X-ray	2.50
<a href="#">1ABW</a>	X-ray	2.50
<a href="#">1ABY</a>	X-ray	2.50
<a href="#">1AJ9</a>	X-ray	2.50
<a href="#">1B86</a>	X-ray	2.50
<a href="#">1BAB</a>	X-ray	2.50
<a href="#">1BBB</a>	X-ray	2.50
<a href="#">1BIJ</a>	X-ray	2.50
<a href="#">1BUW</a>	X-ray	2.50
<a href="#">1BZ0</a>	X-ray	2.50
<a href="#">1BZ1</a>	X-ray	2.50
<a href="#">1BZZ</a>	X-ray	2.50
<a href="#">1C7B</a>	X-ray	2.50
<a href="#">1C7C</a>	X-ray	2.50
<a href="#">1C7D</a>	X-ray	2.50
<a href="#">1CBL</a>	X-ray	2.50
<a href="#">1CBM</a>	X-ray	2.50
<a href="#">1CH4</a>	X-ray	2.50



# Uniprot: contents

Genome annotation databases		Enzyme and pathway databases	
Ensembl	<a href="#">ENST00000335295</a> ; <a href="#">ENSP</a>	Reactome	<a href="#">REACT_604</a> . Hemostasis.
GeneID	<a href="#">3043</a> .	Gene expression databases	
KEGG	<a href="#">hsa:3043</a> .	ArrayExpress	<a href="#">P68871</a> .
UCSC	<a href="#">uc001mae.1</a> human		
Organism-specific		Entry information	
CTD		Entry name	HBB_HUMAN
GeneCards		Accession	Primary (citable) accession number: <b>P68871</b> Secondary accession number(s): <a href="#">A4GX73</a> ↔ <a href="#">Q9UCP9</a>
H-InvDB		Entry history	Integrated into UniProtKB/Swiss-Prot: July 21, 1986 Last sequence update: January 23, 2007 Last modified: June 28, 2011 This is version 97 of the entry and version 2 of the sequence. [ <a href="#">Complete history</a> ]
HGNC		Entry status	Reviewed (UniProtKB/Swiss-Prot)
HPA		Annotation program	<a href="#">Chordata Protein Annotation Program</a> 1 hit.
MIM		<b>Disclaimer</b>	Any medical or genetic information present in this entry is provided for research, e used as a substitute for professional medical advice, diagnosis, treatment or care
neXtProt		Relevant documents	
Orphanet		<a href="#">Human chromosome 11</a> Human chromosome 11: entries, gene names and cross-references to MIM	
PharmGKB		<a href="#">Human entries with polymorphisms or disease mutations</a> List of human entries with polymorphisms or disease mutations	
GenAtlas			
Phylogenomic databases		DrugBank	<a href="#">DB00893</a> . Iron Dextran.
HOVERGEN	<a href="#">HBG009709</a> .	NextBio	<a href="#">12048</a> .
		PMAP-CutDB	<a href="#">P68871</a> .

# Uniprot: search results


UniProtKB

Search
Blast
Align
Retrieve
ID Mapping \*

Search in

Protein Knowledgebase (UniProtKB)

▼

Query

hemoglobin B chain

Search

Advanced Search »

Clear

1 - 25 of 1,243 results for hemoglobin AND B AND chain in UniProtKB sorted by score descending

Browse by

taxonomy

keyword

gene ontology

enzyme class

or pathway

Reduce sequence redundancy to

100%

90%

or 50%

Download

Page 1

of 50

Next »

Results

Customize

Show only reviewed (825) (UniProtKB/Swiss-Prot) or

unreviewed (418) (UniProtKB/TrEMBL) entries

Quote terms: "hemoglobin b"

Restrict term "hemoglobin" to

protein family (24)

gene name (3)

gene ontology (883)

keyword (9)

protein name (970)

web resource (5)

Restrict term "b" to

author (608)

domain (4)

gene name (15)

gene ontology (5)

keyword (4)

protein name (58)

organism (1)

source (7)

strain (5)

taxonomy (6)

tissue (2)

Restrict term "chain" to

author (48)

gene name (2)

gene ontology (7)

keyword (7)

protein name (740)

annotation topic (943)

Restrict term "chain" to

pathway

Accession	Entry name	Status	Protein names	Gene names	Organism	Length
<input type="checkbox"/> <a href="#">P68871</a>	HBB_HUMAN	★	Hemoglobin subunit beta	HBB	Homo sapiens (Human)	147
<input type="checkbox"/> <a href="#">P69905</a>	HBA_HUMAN	★	Hemoglobin subunit alpha	HBA1 HBA2	Homo sapiens (Human)	142
<input type="checkbox"/> <a href="#">P69892</a>	HBG2_HUMAN	★	Hemoglobin subunit gamma-2	HBG2	Homo sapiens (Human)	147
<input type="checkbox"/> <a href="#">P69891</a>	HBG1_HUMAN	★	Hemoglobin subunit gamma-1	HBG1 PRO2979	Homo sapiens (Human)	147
<input type="checkbox"/> <a href="#">P0A592</a>	TRHBN_MYCTU	★	Group 1 truncated hemoglobin glbN	glbN Rv1542c MT1594 MTCY48.23	Mycobacterium tuberculosis	136
<input type="checkbox"/> <a href="#">P02042</a>	HBD_HUMAN	★	Hemoglobin subunit delta	HBD	Homo sapiens (Human)	147

## Search in

## Query

Sequence Clusters (UniRef)

uniprot:(hemoglobin B chain) identity:0.5

Search

Advanced Search »

Clear

1 - 25 of 240 results for uniprot:(hemoglobin AND B AND chain) AND identity:0.5 in UniRef

Browse by taxonomy | Map to UniProtKB or UniParc

Download

Page 1

of 10

## Results Customize

Accession	Status	UniRefCluster name	Size	Members	Organisms	Length	Identity
<input type="checkbox"/> UniRef50_A1B9S1	★	Cluster: TonB-dependent heme/hemoglobin receptor family protein	1	A1B9S1	Paracoccus denitrificans (strain Pd 1222)	686	50%
<input type="checkbox"/> UniRef50_A1U6I3	★	Cluster: TonB-dependent heme/hemoglobin receptor family protein	4	A1U6I3 F2N3G5 C0N4R6 A4XQE8	Marinobacter aquaeolei (strain ATCC 700491 / DSM 11845 / VT8) (Marinobacter hydrocarbonoclasticus (strain DSM 11845)) Pseudomonas stutzeri DSM 4166 Methylophaga thiooxydans DMS010 Pseudomonas mendocina (strain ymp)	657	50%
<input type="checkbox"/> UniRef50_A1WMF8	★	Cluster: TonB-dependent hemoglobin/transferrin/lactoferrin family receptor	1	A1WMF8	Verminephrobacter eiseniae (strain EF01-2)	730	50%
<input type="checkbox"/> UniRef50_A6QC57	★	Cluster: Globin	4	A6QC57 E6WZ29 A6Q1V1 Q30RW0	Sulfurovum sp. (strain NBC37-1) Nitratiraptor salsuginis (strain DSM 16511 / JCM 12458 / E9137-1) Nitratiraptor sp. (strain SB155-2) Sulfurimonas denitrificans (strain ATCC 33889 / DSM 1251) (Thiomicrospira denitrificans (strain ATCC 33889 / DSM 1251))	173	50%
<input type="checkbox"/> UniRef50_A6UBU8	★	Cluster: TonB-dependent hemoglobin/transferrin/lactoferrin family receptor	3	A6UBU8 A9D265 B9JZN4	Sinorhizobium medicae (strain WSM419) (Ensifer medicae) Hoeftia phototrophica DFL-43 Agrobacterium vitis (strain S4 / ATCC BAA-846) (Rhizobium vitis (strain S4))	734	50%

240 results for uniprot:(hemoglobin AND B AND chain) AND identity:0.5 in UniRef

> Download data [compressed](#) or [uncompressed](#)

### Tab-Delimited

Summary information from the result view.

[ [Download](#) | [Open](#) | [Open first 10](#) ]

### Excel

Summary information from the result view for MS Excel™.

[ [Download](#) | [Open](#) | [Open first 10](#) ]

### FASTA

Sequence data in FASTA format.

[ [Download](#) (200 KB\*) | [Open](#) | [Open first 10](#) ]

### XML

Complete data in XML format.

[ [Download](#) (400 KB\*) | [Open](#) | [Open first 10](#) ]

### RDF/XML

Complete data in RDF format.

[ [Download](#) (500 KB\*) | [Open](#) | [Open first 10](#) ]

### List

List of accession numbers.

[ [Download](#) (5 KB\*) | [Open](#) | [Open first 10](#) ]

```
>UniRef50_A1B9S1 TonB-dependent heme/hemoglobin receptor family
MPRHSIRGALLAGTACLTALTFTAPLLAQERAGADSAQSTYVLDQIVLRAGKPKVASEVP
QSVSVVDSRQLEDIAPIHIGEVLATVPGVAGVGSFSGGQGFNIRFGSSGAAASGESGIV
QLIDGEEKYYESYRQGALFVEPDFLRQVEVLRGPGSSSTLYGSGALGGVIAMETIEAGDLI
AEGQTFGGRTKLGYSNPDTVLGSAVALGWRPAEDFEALAAFAWRKLGDTKADGNTTVRA
NSKTPNLLLKAKKTFGDQYVAFSYQHLEAKGDDQDFNQLEGAQVGLFPGFPGWGVGDITT
RDQTARFIWGWNPEDNRYVDLTATLSYTNLTLDVROGDDPDEPIMDSLLGERDYRLWKFK
LSNVADLSGAGYDHLTTGAELVKQDRSSSVSSSHPEAYTRAWAAYALSELTWGDLTIN
SGLRYEKQORTEPKSSVTVTDDTYDADSVPEQVAAIYRLNDSLVSFGSVAFVNRMPVDEL
YDSFMGGAPSGDLKDEKGNIELGLSYHSGILTASDEAVVKLTFRNHIDDMIVRTNAP
APMPAYVNIIDRAYLRGGELEATYSVAAWEFGAASFVNVGVDQDGADLDTLPNNRVTLQAI
WQASDALRLGLRSTLADGRDKPNGTHRAGYGVHDVVFATWVPQGGAAAGIEVHVGVNDVTD
RDYTPATWLSGPAPGRNFKLSVSRSF
>UniRef50_A1U6I3 TonB-dependent heme/hemoglobin receptor family
MANSSPMKQPRFRNTLWLALMAAPLAHAQPVS LDPIQVTADREADADTVVDAETIERF
QADDLEDVFAQPDVSVGGSNSIAQKVYVRGFDPLLNVSDGATQAGALFHHSGRLSVE
PELLKQVEVNAGAGRATEGAGALGGSIRFVTKDPDDLRLPGESAGALVKFGFSNTDGYK
ASGTAFGRLSDNWSTLVSVSQSDHEPFKDGSGDRIAGSDARQQLGFAKLVGQLPADQTIK
LSHEVRTDEGERPQRPQWVVSFNRLYELDGRRDTTTTLNYGYAPAGNALVDLEATVYHTE
SDIEQNVEDRWGRYFGFSRNIGGDLRNTSRFGGHS LTYGVDYREDKVNAGYQEDKRAEQQ
TGEVLGVYLQGDWLWTSRLLSAGARYDDYRLKDNDQRFSEDEVSPNANLAWVVVDGLT
LKAGYAEAFRGPTTQDAFKLEGSENDPDLEGEKARNTVEGFDYRYETFRLSAEVYRSEIK
DAIADPLLPFRESIYKNIGDLES DGYLISAGYQWQALSAGLSFHSNDAEVDGQPLTVYEH
NLLGNTMGDTWIADLAYRWDRNLEFGWQGRFVEGIDNLDTSVGTIDKPGYGVHDLYLHWL
PTGMEDLRLSLTIKNVGDKQYLAHASNADYQHIEDYEGIVGMPEPGRDIRVGLAMRF
```

# Question

**Obtain the sequences of “transcription factors” with less than 50% sequence identity.**

**Find the amino acid sequence of human mitochondrial beta barrel membrane protein VDAC**



# Dataset for transcription factors

UniProtKB > UniProtKB Downloads · Contact · Documentation/Help

Search in **Protein Knowledgebase (UniProtKB)** **transcription factors** Search Clear Fields »

1 - 25 of 10,963 results for **transcription** AND **factors** in UniProtKB sorted by **score** descending

Browse by taxonomy, keyword, gene ontology, enzyme class or pathway | Reduce sequence redundancy to 100%, 90% or 50% | Customize display **Download...**

› Show only **reviewed** ★ (UniProtKB/Swiss-Prot) or **unreviewed** ★ (UniProtKB/TrEMBL) entries

› Quote terms: "transcription factors"

› Restrict term "transcription" to **protein family**, **gene ontology**, **keyword**, **protein name**

› Restrict term "factors" to **protein family**, **gene ontology**, **protein name**, **web resource**

Page 1 of 439 | Next »

All	Accession	Entry name	Status	Protein names	Gene names	Organism	Length
<input checked="" type="checkbox"/>	<a href="#">P0A4H2</a>	BVGA_BORPE	★	<b>Virulence factors putative positive transcription regulator bvgA</b>	<b>bvgA</b> (BP1878)	Bordetella pertussis	209
<input checked="" type="checkbox"/>	<a href="#">P0A4H3</a>	BVGA_BORBR	★	<b>Virulence factors putative positive transcription regulator bvgA</b>	<b>bvgA</b> (BB2994)	Bordetella bronchiseptica (Alcaligenes bronchisepticus)	209
<input checked="" type="checkbox"/>	<a href="#">P0A4H4</a>	BVGA_BORPA	★	<b>Virulence factors putative positive transcription regulator bvgA</b>	<b>bvgA</b> (BPP3028)	Bordetella parapertussis	209
<input checked="" type="checkbox"/>	<a href="#">P18850</a>	ATF6A_HUMAN	★	<b>Cyclic AMP-dependent transcription factor ATF-6 alpha</b> (cAMP-dependent transcription factor ATF-6 alpha) (Activating transcription factor 6 alpha) (ATF6-alpha) [Cleaved into: Processed cyclic AMP-dependent transcription factor ATF-6 alpha]	<b>ATF6</b>	Homo sapiens (Human)	670
<input checked="" type="checkbox"/>	<a href="#">A6H8Y1</a>	BDP1_HUMAN	★	<b>Transcription factor TFIIIB component B'' homolog</b> (Transcription factor IIIB 150) (TFIIIB150) (Transcription factor-like nuclear regulator)	<b>BDP1</b> (KIAA1241) (KIAA1689) (TFNR)	Homo sapiens (Human)	2,624
<input checked="" type="checkbox"/>	<a href="#">Q04636</a>	POB3_YEAST	★	<b>FACT complex subunit POB3</b> (Facilitates chromatin transcription complex subunit POB3)	<b>POB3</b> (YML069W)	Saccharomyces cerevisiae (Baker's yeast)	552

>sp|P18485|1A12\_SOLLC 1-aminocyclopropane-1-carboxylate synthase 2 OS=Solanum lycopersicum GN=ACS2 PE=1 SV=2  
MGFEIAKTNLSILSKLATNEEHGENSPYFDGWKAYDSDPFHPLKNPNGVIQMGLAENQLCL  
DLIEDWIKRNPKGSIKSEGIKSFKAIFQDYHGLPEFRKAIKFMKTRGGRVRFDPER  
VVMAGGATGANETIIFCLADPGDAFLVPSPIYPFNRDLRWRTGVQLIPIHCESNNFKI  
TSKAVKEAYENAQKSNIKVKGLILTNPSNPLGTTLDKDTLKSVLSTNQHNHILVCDIY  
AATVFDTPQFVSIAEILDEQEMTYCNKDLVHIVYSLSKDMGLPGFRVGIISFNDDVUNC  
ARKMSSFGVLSTQTQYFLAAMLSDKFDNLFRESAMRLGKRHKHFTNGLEVVGKICLKN  
NAGLFCWMDLRPLLRETFDSEMSLWRVVIINDVKLVNVPSSFECPGWFVCFANMDD  
GTVDIALARIRRFVGVKSGDKSSMEKKQWKKNNLRLSFSKRMVDESVLSPSSPIPP  
SPLVR

>sp|P16375|7UP1\_DROME Steroid receptor seven-up, isoforms B/C OS=Drosophila melanogaster GN=svp PE=1 SV=1  
MCASPSTAPGFFNPRPQSGAELSAFDIGLSRSMGLGVPPHSAWHEPPASLGGLHAASAG  
PGTTTGSVATGGGGTTPSSVASQSSAVIKQDLSCPSLNQAGSGHHPGIKEDLSSSLPSAN  
GGSAGGHHSGSGSGSGSVNPGHSDMLPLIKGHGQDMLTSIKGQPTGCGSTTPSSQANS  
SHSQSSNSGSQIDSKQNICVVCVCGDKSSGKHGQFTCEGCKSFFKRSVRRNLTYSCRGS  
NCPIDQHHRNQCCYCRLLKCKLMGMRREAVQGRVPPTQPLAGMHGQYQIANGDPMGIA  
GFNGHSSYSSYISLLLRAEYPYTSRYGQCMQPNINMGIDNICELAAARLLFSAVEWAKNIP  
FFPELQVTDQVALLRLVWSELVFLNASQCSMPLHVAPLLAAAGLHASPMAADRVVAFMDH  
IRIFQEVEKLKALHVDSEAESCLKAIVLFTTDACGLSDVTHIESLQEKSQLALEEYCRT  
QYPNQPTRFGKLLRLPLSLRTVSSQVIEQLFFVRLVGKTP IETLIRDMLLSGNSFSWPYL  
PSM

>sp|P16376|7UP2\_DROME Steroid receptor seven-up, isoform A OS=Drosophila melanogaster GN=svp PE=2 SV=3  
MCASPSTAPGFFNPRPQSGAELSAFDIGLSRSMGLGVPPHSAWHEPPASLGGLHAASAG  
PGTTTGSVATGGGGTTPSSVASQSSAVIKQDLSCPSLNQAGSGHHPGIKEDLSSSLPSAN  
GGSAGGHHSGSGSGSGSVNPGHSDMLPLIKGHGQDMLTSIKGQPTGCGSTTPSSQANS  
SHSQSSNSGSQIDSKQNICVVCVCGDKSSGKHGQFTCEGCKSFFKRSVRRNLTYSCRGS  
NCPIDQHHRNQCCYCRLLKCKLMGMRREAVQGRVPPTQPLAGMHGQYQIANGDPMGIA  
GFNGHSSYSSYISLLLRAEYPYTSRYGQCMQPNINMGIDNICELAAARLLFSAVEWAKNIP  
FFPELQVTDQVALLRLVWSELVFLNASQCSMPLHVAPLLAAAGLHASPMAADRVVAFMDH  
IRIFQEVEKLKALHVDSEAESCLKAIVLFTTGKLLDILYKDVPAALLTKVSALLGKGSTA  
SNDDVLAVVRDHLDELNRQEQESQAQQQAAPLHLAAFMNCVAGVEAAVQQAEEQAQVPTSSA  
SASVSAPLVPSAGSAFSSCQAKSAGSEMDLLASLYAQQAATPPSSGGGDASGHNNSSGLG  
ASLPTQSQSGSSSRNLASPLSTSLATAPAPASASAPAPVPTSSVAQVPVPAPVPTSSA  
SSSSLGGGAYQTPSAAAAAAAMFHYQTPPRAAFGSAFDMFHHSTPFGVGVGHAHALAHSS  
GSGSASFSGSPSYRSPYSLAGSRWQL

## Search in

Protein Knowledgebase (UniProtKB) ▼

## Query

mitochondrial beta barrel membrane protein  
AND human and VDAC

Search

Advanced Search »

Clear

3 results for **mitochondrial** AND **beta** AND **barrel** AND **membrane** AND **protein** AND **human** AND **VDAC** in UniProtKB sorted by **score** descendingBrowse by [taxonomy](#), [keyword](#), [gene ontology](#), [enzyme class](#) or [pathway](#) | Reduce sequence redundancy to [100%](#), [90%](#) or [50%](#) |

Download



Page 1 of 1

Results [Customize](#)

- > Restrict term "mitochondrion" to [protein family](#) (3), [gene ontology](#) (3), [keyword](#) (3), [location](#) (3), [protein name](#) (3)
- > Restrict term "beta" to [keyword](#) (3), [annotation topic](#) (1)
- > Restrict term "membrane" to [gene ontology](#) (3), [keyword](#) (3), [location](#) (3), [protein name](#) (3)
- > Restrict term "protein" to [gene ontology](#) (3), [keyword](#) (2), [protein name](#) (3)
- > Restrict term "mitochondrion" to [pathway](#)
- > Restrict term "beta" to [pathway](#)
- > Restrict term "membrane" to [pathway](#)
- > Restrict term "protein" to [pathway](#)

	Accession	Entry name	Status	Protein names	Gene names	Organism	Length
<input type="checkbox"/>	<a href="#">P21796</a>	VDAC1_HUMAN	★	Voltage-dependent anion-selective channel pro...	VDAC1 VDAC	Homo sapiens (Human)	283
<input type="checkbox"/>	<a href="#">P45880</a>	VDAC2_HUMAN	★	Voltage-dependent anion-selective channel pro...	VDAC2	Homo sapiens (Human)	294
<input type="checkbox"/>	<a href="#">Q9Y277</a>	VDAC3_HUMAN	★	Voltage-dependent anion-selective channel pro...	VDAC3	Homo sapiens (Human)	283

## Sequences

Sequence	Length	Mass (Da)	Tools
<input type="checkbox"/> P21796 [UniParc]. <a href="#">FASTA</a>	283	30,773	Blast <input type="button" value="go"/>

Last modified January 23, 2007. Version 2.  
Checksum: 89BA3378B04020D5

```

      10      20      30      40      50      60
MAVPPTYADL GKSARDVFTK GYGFGGLIKLD LKTKSENGLE FTSSGSANTE TTKVTGSLET

      70      80      90     100     110     120
KYRWTEYGLT FTEKWNTDNT LGTEITVEDQ LARGLKLTFD SSFSPNTGKK NAKIKTGYKR

     130     140     150     160     170     180
EHINLGCDMD FDIAGPSIRG ALVLGYEGWL AGYQMNFFETA KSRVTQSNFA VGYKTDEFQL

     190     200     210     220     230     240
HTNVNDGTEF GGSIIYQKVNK KLETAVNLAW TAGNSNTRFG IAAKYQIDPD ACFSAKVNNS

     250     260     270     280
SLIGLGYTQT LKPGIKLTLA ALLDGKNVNA GGHLGLGLE FQA

```

```

>sp|P21796|VDAC1_HUMAN Voltage-dependent anion-selective channel protein 1 OS=Homo sapiens GN=VDAC1 PE=1 SV=2
MAVPPTYADL GKSARDVFTK GYGFGGLIKLD LKTKSENGLE FTSSGSANTE TTKVTGSLET
KYRWTEYGLT FTEKWNTDNT LGTEITVEDQ LARGLKLTFD SSFSPNTGKK NAKIKTGYKR
EHINLGCDMD FDIAGPSIRG ALVLGYEGWL AGYQMNFFETA KSRVTQSNFA VGYKTDEFQL
HTNVNDGTEF GGSIIYQKVNK KLETAVNLAW TAGNSNTRFG IAAKYQIDPD ACFSAKVNNS
SLIGLGYTQT LKPGIKLTLA ALLDGKNVNA GGHLGLGLE FQA

```