1

Introduction

The endeavour to write a textbook on the Structure and Function of Human Blood Plasma Proteins bears an inherent risk regarding the selection of the proteins to be discussed, because this selection will always be ambiguous depending on the applied definition of the term blood plasma protein. Based on the classification proposed in the 1970s by Frank W. Putnam in his famous book series The Plasma Proteins: Structure, Function and Genetic Control, in 2002 N. Leigh Anderson and Norman G. Anderson elaborated on the classification of human blood plasma proteins in a publication in Molecular and Cellular Proteomics, ‘The human plasma proteome: history, character and diagnostic prospects’ (for details see Chapter 3). The classification proposed by Anderson and Anderson clearly shows that human plasma contains the most comprehensive version of the human proteome. The complexity of the ‘plasma proteome’ is quickly understood when all the various forms of blood plasma proteins present in plasma are considered: precursor and mature forms, splice variants, degradation products and, of course, all combinations of posttranslational modifications. Of course, the scope and the available space in this book require a rigid selection of the blood plasma proteins discussed in some detail. The selection is primarily based on the classification introduced by Anderson and Anderson and on personal considerations of the authors.

The main sources of information used in this book including the references therein were the following:

1. Databases
   \begin{itemize}
   \item UniProt Knowledgebase (Swiss-Prot and TrEMBL)
   \item PROSITE: database of protein domains, families and functional sites
   \item Protein Data Bank (PDB): an information portal to biological macromolecular structures
   \item MIM: Mendelian Inheritance in Man (at the NCBI)
   \end{itemize}

2. Books
   \begin{itemize}
   \item Human Protein Data (Haeberli, 1998)
   \item Biochemical Pathways (Michel, 1998)
   \item Introduction to Protein Structure (Branden and Tooze, 1999)
   \item Biochemistry (Voet and Voet, 2004)
   \item Proteins: Structure and Function (Whitford, 2005)
   \end{itemize}

3. Articles
   \begin{itemize}
   \item ‘The human plasma proteome’ (Anderson and Anderson, 2002)
   \item ‘The human plasma proteome’ (Anderson et al., 2004)
   \item ‘The human serum proteome’ (Pieper et al., 2003)
   \item ‘Exploring the human plasma proteome’ (Omenn, 2005)
   \end{itemize}

This book is divided into three parts:

\begin{itemize}
\item Part I
  Blood Components
  Blood Plasma Proteins
\end{itemize}
If not otherwise stated, the proteins discussed in this book are of human origin. The used protein name is the main name in Swiss-Prot and in some cases common synonyms are also given. All shown three-dimensional (3D) structures are from the Protein Data Bank (PDB). No model structures were included and the structures were either determined by X-ray diffraction or by nuclear magnetic resonance (NMR) spectroscopy. In a few cases nonhuman 3D structures are presented if the corresponding human 3D structure has not yet been determined. For enzymes the common EC classification system is given. Many domains, motifs and repeats or a certain stretch of sequence in a protein are characterised by a typical signature. In this book the signatures of the PROSITE database are used. In many cases diseases related to a certain protein are briefly mentioned and references to the disease database MIM (Mendelian Inheritance in Man) are given.

A limited number of references is given in each chapter and in the Data Sheets. Special emphasis was put on the quality of the references and the journals and their worldwide availability.

The information of each protein discussed in this book in some detail is summarised at the end of each chapter in a Data Sheet, where the most important data of each protein can be found at a glance. Proteins mentioned in the text but not discussed are compiled in the Appendix (Table A.1, human and Table A.2, nonhuman, with the corresponding reference to the database entry).

Each Data Sheet is divided into four sections:

Fact Sheet  
Description  
Structure  
Biological Function

<table>
<thead>
<tr>
<th>Table 1.1</th>
<th>Useful universal resource locators (URL).</th>
</tr>
</thead>
<tbody>
<tr>
<td>URL</td>
<td>Description</td>
</tr>
<tr>
<td><a href="http://www.expasy.org">www.expasy.org</a></td>
<td>ExPASy (Expert Protein Analysis System): proteomics server of the Swiss Institute of Bioinformatics (SIB) dedicated to protein analysis</td>
</tr>
<tr>
<td></td>
<td>Databases: UniProt (SwissProt + TrEMBL), PROSITE, SWISS-MODEL Proteomics and Sequence Analysis Tools</td>
</tr>
<tr>
<td>pir.georgetown.edu</td>
<td>PIR (Protein Information Resource): integrated protein information resource for genomic and proteomic research</td>
</tr>
<tr>
<td></td>
<td>Databases: PIRSF, iProClass, iProLink</td>
</tr>
<tr>
<td><a href="http://www.rcsb.org/pdb">www.rcsb.org/pdb</a></td>
<td>RCSB (Research Collaboratory of Structural Bioinformatics): information portal to biological macromolecular structures</td>
</tr>
<tr>
<td></td>
<td>Database: PDB</td>
</tr>
<tr>
<td><a href="http://www.ebi.ac.uk">www.ebi.ac.uk</a></td>
<td>EBI (European Bioinformatics Institute): freely available data and bioinformatic services</td>
</tr>
</tbody>
</table>
The Fact Sheet contains the following data:

- **Classification:** Swiss-Prot
- **Abbreviations:** Most common
- **Structures/motifs:** Swiss-Prot and PROSITE
- **DB/PDB entries:** Swiss-Prot/Protein Data Bank
- **MW/length:** Mature Protein (without PTMs)
- **Concentration:** Approximate or range (if available), adult
- **Half-life:** Approximate or range (if available)
- **PTMs:** Swiss-Prot (most common)
- **References:** Key reference(s) on sequence and 3D structure

In the section Description each protein is briefly described (usually by only one sentence).
In the section Structure the main structural features of a protein are briefly summarised and, if available, a short description of the 3D structure is given. If available, a figure of the 3D structure is included.
Finally, in the section Biological Function the main, usually physiological, functions are briefly summarised.

A limited number of useful universal resource locators (URLs) containing reliable protein data from curated and annotated databases are tabulated in Table 1.1.

REFERENCES
