FOURTH EDITION

HANDBOOK OF BIOCHEMISTRY AND MOLECULAR BIOLOGY

EDITED BY
ROGER L. LUNDBLAD
FIONA M. MACDONALD

CRC Press
Taylor & Francis Group

FOURTH EDITION

HANDBOOK OF BIOCHEMISTRY AND MOLECULAR BIOLOGY

EDITED BY
ROGER L. LUNDBLAD
FIONA M. MACDONALD

CRC Press
Taylor & Francis Group
Boca Raton London New York

CRC Press is an imprint of the
Taylor & Francis Group, an **informa** business

CRC Press
Taylor & Francis Group
6000 Broken Sound Parkway NW, Suite 300
Boca Raton, FL 33487-2742

Printed in the United States of America on acid-free paper
10 9 8 7 6 5 4 3 2 1

International Standard Book Number: 978-0-8493-9168-2 (Hardback)

Library of Congress Cataloging-in-Publication Data

Handbook of biochemistry and molecular biology / edited by Roger L. Lundblad and Fiona M. Macdonald.
p. ; cm.
Includes bibliographical references and index.
ISBN-13: 978-0-8493-9168-2 (alk. paper)
ISBN-10: 0-8493-9168-7 (alk. paper)
1. Biochemistry--Handbooks, manuals, etc. 2. Molecular biology--Handbooks, manuals, etc. I. Lundblad, Roger L. II. Macdonald, Fiona.
[DNLM: 1. Biochemical Phenomena--Tables. QU 16 H2363 2010]

QH345.H347 2010
572--dc22 2009042039

Visit the Taylor & Francis Web site at
http://www.taylorandfrancis.com

and the CRC Press Web site at
http://www.crcpress.com

FOURTH EDITION

HANDBOOK OF BIOCHEMISTRY AND MOLECULAR BIOLOGY

This work is dedicated to the many scientists of the "The Greatest Generation" who contributed to the base of our knowledge of biochemistry and molecular biology.

Roger L. Lundblad, Ph.D.

To my parents, Pat and Paul Macdonald, whose unwavering love and support has been my guiding light.

Fiona M. Macdonald, Ph.D., F.R.S.C.

FOREWORD

For almost a century, CRC Press has been a leader in providing concise compilations of scientific data for researchers, teachers, and students. The *CRC Handbook of Chemistry and Physics,* which first appeared in 1913 and is now in its 90th edition, is the basic source of physical science data that most chemists, physicists, and engineers turn to. Other widely used handbooks from CRC Press cover materials science, engineering, and mathematics. Many specialized handbooks have also appeared under the CRC imprint, ranging from semiconductors to lipids.

The 1968 publication of the *CRC Handbook of Biochemistry,* edited by Herbert A. Sober, marked a milestone in bioscience data. Appearing just 15 years after Watson and Crick elucidated the structure of DNA, the subtitle of this work, *Selected Data for Molecular Biology,* was a recognition that molecular biology was the new frontier of the biosciences. This was followed by the multivolume *Handbook of Biochemistry and Molecular Biology,* edited by Gerald D. Fasman, and its single volume abridged version *Practical Handbook of Biochemistry and Molecular Biology,* which appeared in 1989. The intervening 20 years has seen an explosion of data in this field and an exponential growth in the translation of recent discoveries into new technology. This new *Handbook of Biochemistry,* edited by Roger Lundblad and Fiona M. Macdonald, is thus a welcome addition to the list of handbooks available through CRC Press. I am sure it will find heavy use in both basic research and biotechnology.

David R. Lide
Editor-in-Chief, CRC Handbook of Chemistry and Physics

PREFACE

This is the fourth edition of the *Handbook of Biochemistry and Molecular Biology*. The first edition was published as a single volume in 1968 under the guidance of Herbert Sober. The second edition appeared in 1970 and the third, with Gerald Fasman as editor, appeared in eight volumes published in 1975–6. This increase in size reflected the rapid advances in knowledge in the then relatively new field of molecular biology.

It is intended that current *Handbook of Biochemistry and Molecular Biology* be a companion volume to the *CRC Handbook of Chemistry and Physics*—a single volume ready-reference work that will find a home on the bookshelves of biochemists and molecular biologists everywhere.

This fourth edition contains materials from the previous editions as well as extensive new material. Staying within the confines of a single volume has meant difficult decisions on which tables to include, and the editors welcome feedback from readers. The advent of electronic media allows for more frequent updating and it is hoped that any infelicities in our selection may be readily rectified. Additionally, suggestions on new topics for this Handbook and notification of errors are always appreciated. Address all comments to Editor, *Handbook of Biochemistry and Molecular Biology*, Taylor & Francis Group, 6000 Broken Sound Parkway NW, Suite 300, Boca Raton, FL 33487.

Much of the current content is derived from the research of the giants of biochemistry and molecular biology in the three decades following World War II. While it seems the vogue to develop new names and descriptions for old, established concepts, biochemistry and molecular biology continue to be the mainstays of current biomedical research.

Roger L. Lundblad
lundbladr@bellsouth.net

Fiona M. Macdonald
fiona.macdonald@taylorandfrancis.com
March 2010

ACKNOWLEDGMENTS

This work would not have been possible without the outstanding support of Jill Jurgensen and Glen Butler of Taylor & Francis. The support of Professor Edward A. Dennis is acknowledged. The help of various research librarians at the University of North Carolina at Chapel Hill is also acknowledged. Professors Charles Craik of the University of California at San Francisco and Bryce Plapp of the University of Iowa provided advice on selection of materials to be included. However, the editors take all responsibility for the selection of content.

EDITORS

Roger L. Lundblad, PhD

Roger L. Lundblad is a native of San Francisco, California. He received his undergraduate education at Pacific Lutheran University and his PhD degree in biochemistry at the University of Washington. After postdoctoral work in the laboratories of Stanford Moore and William Stein at the Rockefeller University, he joined the faculty of the University of North Carolina at Chapel Hill. He joined the Hyland Division of Baxter Healthcare in 1990. Currently Dr. Lundblad is an independent consultant and writer in biotechnology in Chapel Hill, North Carolina. He is an adjunct professor of Pathology at the University of North Carolina at Chapel Hill and editor-in-chief of the Internet Journal of Genomics and Proteomics.

Fiona M. Macdonald, PhD, F.R.S.C.

Fiona M. Macdonald received her BSc in chemistry from Durham University, UK. She obtained her PhD in inorganic biochemistry at Birkbeck College, University of London, studying under Peter Sadler. Having spent most of her career in scientific publishing, she is now at Taylor & Francis and is involved in developing chemical information products.

TABLE OF CONTENTS

Section I
Amino Acids, Peptides, and Proteins

PROPERTIES OF AMINO ACIDS

This table gives selected properties of some important amino acids and closely related compounds. The first part of the table lists the 20 "standard" amino acids that are the basic constituents of proteins. The second part includes other amino acids and related compounds of biochemical importance. Within each part of the table the compounds are listed by name in alphabetical order. Structures are given in the following table.

Symbol : Three-letter symbol for the standard amino acids
M_r : Molecular weight
t_m: Melting point
pK_a, pK_b, pK_c, pK_d : Negative of the logarithm of the acid dissociation constants for the COOH and NH_2 groups (and, in some cases, other groups) in the molecule (at 25°C)
pI: pH at the isoelectric point
S: Solubility in water in units of grams of compound per kilogram of water; a temperature of 25°C is understood unless otherwise stated in a superscript. When quantitative data are not available, the notations sl.s. (for slightly soluble), s. (for soluble), and v.s. (for very soluble) are used.
V_2^0: Partial molar volume in aqueous solution at infinite dilution (at 25°C)

Data on the enthalpy of formation of many of these compounds are included in the table "Heat of Combustion, Enthalpy and Free Energy of Formation of Amino Acids and Related Compounds" on p. 69 of this Handbook. Absorption spectra and optical rotation data can be found in Reference 3. Partial molar volume is taken from Reference 5; other thermodynamic properties, including solubility as a function of temperature, are given in References 3 and 5. Most of the pK values come from References 1, 6, and 7.

References

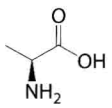
1. Dawson, R. M. C., Elliott, D. C., Elliott, W. H., and Jones, K. M., *Data for Biochemical Research*, Third Edition, Clarendon Press, Oxford, 1986.
2. O'Neil, Maryadele J., Ed., *The Merck Index, Fourteenth Edition*, Merck & Co., Rahway, NJ, 2006.
3. Sober, H. A., Ed., *CRC Handbook of Biochemistry. Selected Data for Molecular Biology*, CRC Press, Boca Raton, FL, 1968.
4. Voet, D., and Voet, J. G., *Biochemistry, Second Edition*, John Wiley & Sons, New York, 1995.
5. Hinz, H. J., Ed., *Thermodynamic Data for Biochemistry and Biotechnology*, Springer–Verlag, Heidelberg, 1986.
6. Fasman, G. D., Ed. *Practical Handbook of Biochemistry and Molecular Biology*, CRC Press, Boca Raton, FL, 1989.
7. Smith, R. M., and Martell, A. E., *NIST Standard Reference Database 46: Critically Selected Stability Constants of Metal Complexes Database*, Version 3.0, National Institute of Standards and Technology, Gaithersburg, MD, 1997.
8. Ramasami, P., *J. Chem. Eng. Data*, 47, 1164, 2002.

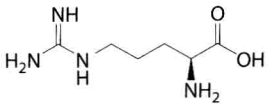
Symbol	Name	Mol. Form	M_r	t_m/°C	pK_a	pK_b	pK_c	pK_d	pI	S/g kg^{-1}	V_2^0/cm^3 mol^{-1}
Ala	L-Alanine	$C_3H_7NO_2$	89.09	297	2.33	9.71			6.00	166.9	60.54
Arg	L-Arginine	$C_6H_{14}N_4O_2$	174.20	244	2.03	9.00	12.10		10.76	182.6	127.42
Asn	L-Asparagine	$C_4H_8N_2O_3$	132.12	235	2.16	8.73			5.41	25.1	78.0
Asp	L-Aspartic acid	$C_4H_7NO_4$	133.10	270	1.95	9.66	3.71		2.77	5.04	74.8
Cys	L-Cysteine	$C_3H_7NO_2S$	121.16	240	1.91	10.28	8.14		5.07	v.s.	73.45
Gln	L-Glutamine	$C_5H_{10}N_2O_3$	146.14	185	2.18	9.00			5.65	42.5	
Glu	L-Glutamic acid	$C_5H_9NO_4$	147.13	160	2.16	9.58	4.15		3.22	8.6	89.85
Gly	Glycine	$C_2H_5NO_2$	75.07	290	2.34	9.58			5.97	250.2	43.26
His	L-Histidine	$C_6H_9N_3O_2$	155.15	287	1.70	9.09	6.04		7.59	43.5	98.3
Ile	L-Isoleucine	$C_6H_{13}NO_2$	131.17	284	2.26	9.60			6.02	34.2	105.80
Leu	L-Leucine	$C_6H_{13}NO_2$	131.17	293	2.32	9.58			5.98	22.0	107.77
Lys	L-Lysine	$C_6H_{14}N_2O_2$	146.19	224	2.15	9.16	10.67		9.74	5.8	108.5
Met	L-Methionine	$C_5H_{11}NO_2S$	149.21	281	2.16	9.08			5.74	56	105.57
Phe	L-Phenylalanine	$C_9H_{11}NO_2$	165.19	283	2.18	9.09			5.48	27.9	121.5
Pro	L-Proline	$C_5H_9NO_2$	115.13	221	1.95	10.47			6.30	1622	82.76
Ser	L-Serine	$C_3H_7NO_3$	105.09	228	2.13	9.05			5.68	250	60.62
Thr	L-Threonine	$C_4H_9NO_3$	119.12	256	2.20	8.96			5.60	98.1	76.90
Trp	L-Tryptophan	$C_{11}H_{12}N_2O_2$	204.23	289	2.38	9.34			5.89	13.2	143.8
Tyr	L-Tyrosine	$C_9H_{11}NO_3$	181.19	343	2.24	9.04	10.10		5.66	0.46	
Val	L-Valine	$C_5H_{11}NO_2$	117.15	315	2.27	9.52			5.96	88	90.75
	N-Acetylglutamic acid	$C_7H_{11}NO_5$	189.17	199						s.	
	N^6-Acetyl-L-lysine	$C_8H_{16}N_2O_3$	188.22	265	2.12	9.51					
	β-Alanine	$C_3H_7NO_2$	89.09	200	3.51	10.08				723.6	58.28

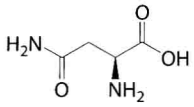
PROPERTIES OF AMINO ACIDS (Continued)

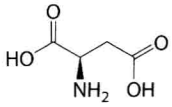
Name	Mol. Form	M_r	t_m/°C	pK_a	pK_b	pK_c	pK_d	pI	S/g kg^{-1}	V_2^0/cm^3 mol^{-1}
2-Aminoadipic acid	$C_6H_{11}NO_4$	161.16	207	2.14	4.21	9.77		3.18	2.2[40]	
DL-2-Aminobutanoic acid	$C_4H_9NO_2$	103.12	304	2.30	9.63			6.06	210	75.6
DL-3-Aminobutanoic acid	$C_4H_9NO_2$	103.12	194.3	3.43	10.05			7.30	1250	76.3
4-Aminobutanoic acid	$C_4H_9NO_2$	103.12	203	4.02	10.35				971	73.2
10-Aminodecanoic acid	$C_{10}H_{21}NO_2$	187.28	188.5							167.3
7-Aminoheptanoic acid	$C_7H_{15}NO_2$	145.20	195						v.s.	120.0
6-Aminohexanoic acid	$C_6H_{13}NO_2$	131.17	205					7.29	863	104.2
L-3-Amino-2-methylpropanoic acid	$C_4H_9NO_2$	103.12	185						s.	
2-Amino-2-methylpropanoic acid	$C_4H_9NO_2$	103.12	335	2.36	10.21			5.72	137	77.55
9-Aminononanoic acid	$C_9H_{19}NO_2$	173.26	191							151.3
8-Aminooctanoic acid	$C_8H_{17}NO_2$	159.23	192							136.1
5-Amino-4-oxopentanoic acid	$C_5H_9NO_3$	131.13	118	4.05	8.90					
5-Aminopentanoic acid	$C_5H_{11}NO_2$	117.15	157 dec						s.	87.6
o-Anthranilic acid	$C_7H_7NO_2$	137.14	146	2.05	4.95				3.5[14]	
Azaserine	$C_5H_7N_3O_4$	173.13	150		8.55				v.s.	
Canavanine	$C_5H_{12}N_4O_3$	176.17	172	2.50	6.60	9.25		7.93	v.s.	
L-γ-Carboxyglutamic acid	$C_6H_9NO_6$	191.14	167	1.70	9.90	4.75	3.20			
Carnosine	$C_9H_{14}N_4O_3$	226.23	260	2.51	9.35	6.76			322	
Citrulline	$C_6H_{13}N_3O_3$	175.19	222	2.32	9.30			5.92	s.	
Creatine	$C_4H_9N_3O_2$	131.13	303	2.63	14.30				16	
L-Cysteic acid	$C_3H_7NO_5S$	169.16	260	1.89	8.70	1.30			v.s.	
L-Cystine	$C_6H_{12}N_2O_4S_2$	240.30	260	1.50	8.80	2.05	8.03		0.11	
2,4-Diaminobutanoic acid	$C_4H_{10}N_2O_2$	118.13	118.1	1.85	8.24	10.44		9.27	s.	
3,5-Dibromo-L-tyrosine	$C_9H_9Br_2NO_3$	338.98	245						2.72	
3,5-Dichloro-L-tyrosine	$C_9H_9Cl_2NO_3$	250.08	247						1.97	
3,5-Diiodo-L-tyrosine	$C_9H_9I_2NO_3$	432.98	213	2.12	9.10	6.16			0.62	
Dopamine	$C_8H_{11}NO_2$	153.18			10.36	8.88			s.	
L-Ethionine	$C_6H_{13}NO_2S$	163.24	273	2.18	9.05	13.10				
N-Glycylglycine	$C_4H_8N_2O_3$	132.12	263	3.13	8.10				225	
Guanidinoacetic acid	$C_3H_7N_3O_2$	117.11	282	2.82					5	
Histamine	$C_5H_9N_3$	111.15	83		9.83	6.11			v.s.	
L-Homocysteine	$C_4H_9NO_2S$	135.19	232	2.15	8.57	10.38		5.55	s.	
Homocystine	$C_8H_{16}N_2O_4S_2$	268.35	264	1.59	9.44	2.54	8.52		0.2	
L-Homoserine	$C_4H_9NO_3$	119.12	203	2.27	9.28			6.17	1100	
3-Hydroxy-DL-glutamic acid	$C_5H_9NO_5$	163.13	209					3.28		
5-Hydroxylysine	$C_6H_{14}N_2O_3$	162.19		2.13	8.85	9.83		9.15		
trans-4-Hydroxy-L-proline	$C_5H_9NO_3$	131.13	274	1.82	9.47			5.74	361	84.49
L-3-Iodotyrosine	$C_9H_{10}INO_3$	307.08	205	2.20	9.10	8.70			sl.s.	
L-Kynurenine	$C_{10}H_{12}N_2O_3$	208.21	194						sl.s.	
L-Lanthionine	$C_6H_{12}N_2O_4S$	208.24	294						1.5	
Levodopa	$C_9H_{11}NO_4$	197.19	277	2.32	8.72	9.96	11.79		5[20]	
L-1-Methylhistidine	$C_7H_{11}N_3O_2$	169.18	249	1.69	8.85	6.48			200	
L-Norleucine	$C_6H_{13}NO_2$	131.17	301	2.31	9.68			6.09	15	107.7
L-Norvaline	$C_5H_{11}NO_2$	117.15	307	2.31	9.65				107	91.8
L-Ornithine	$C_5H_{12}N_2O_2$	132.16	140	1.94	8.78	10.52		9.73	v.s.	
O-Phosphoserine	$C_3H_8NO_6P$	185.07	166	2.14	9.80	5.70				
L-Pyroglutamic acid	$C_5H_7NO_3$	129.12	162	3.32						
Sarcosine	$C_3H_7NO_2$	89.09	212	2.18	9.97				428	
Taurine	$C_2H_7NO_3S$	125.15	328	-0.3	9.06				105	
L-Thyroxine	$C_{15}H_{11}I_4NO_4$	776.87	235	2.20	10.01	6.45			sl.s.	

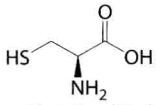
STRUCTURES OF COMMON AMINO ACIDS

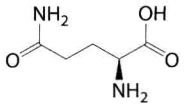

L-Alanine (Ala)

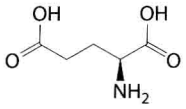

L-Arginine (Arg)

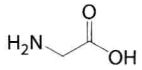

L-Asparagine (Asn)

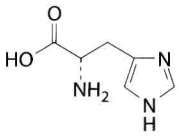

L-Aspartic acid (Asp)

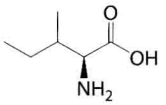

L-Cysteine (Cys)

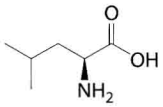

L-Glutamine (Gln)

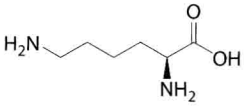

L-Glutamic acid (Glu)

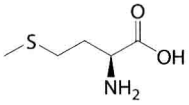

Glycine (Gly)

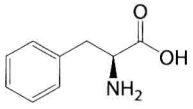

L-Histidine (His)

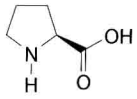

L-Isoleucine (Ile)

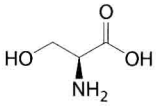

L-Leucine (Leu)

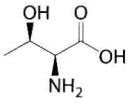

L-Lysine (Lys)

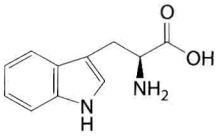

L-Methionine (Met)

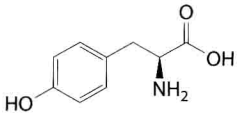

L-Phenylalanine (Phe)

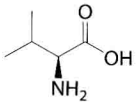

L-Proline (Pro)

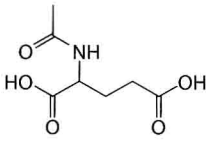

L-Serine (Ser)

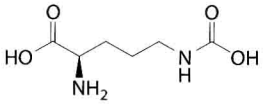

L-Threonine (Thr)

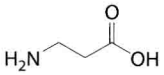

L-Tryptophan (Trp)

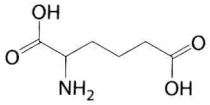

L-Tyrosine (Tyr)

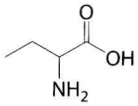

L-Valine (Val)

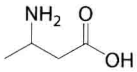

N-Acetylglutamic acid

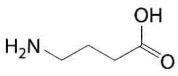

N^6-Acetyl-L-lysine

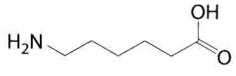

β-Alanine

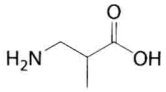

2-Aminoadipic acid

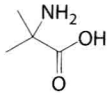

DL-2-Aminobutanoic acid

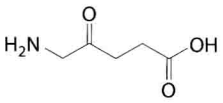

DL-3-Aminobutanoic acid

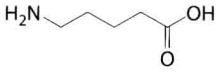

4-Aminobutanoic acid

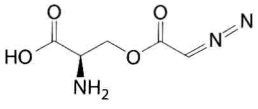

6-Aminohexanoic acid

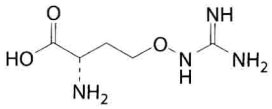

L-3-Amino-2-methylpropanoic acid

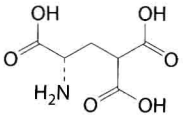

2-Amino-2-methylpropanoic acid

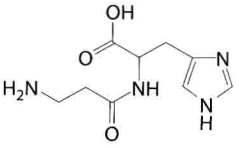

5-Amino-4-oxopentanoic acid


5-Aminopentanoic acid


Azaserine


Canavanine


L-γ-Carboxyglutamic acid


Carnosine