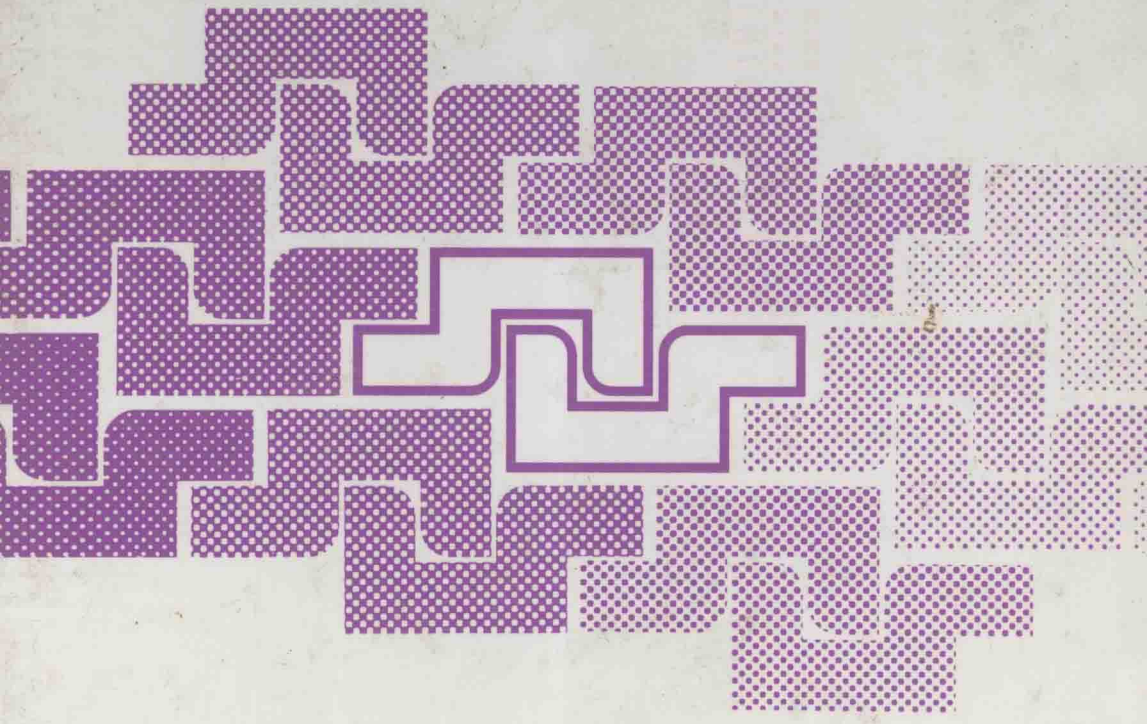

New Methods for the Study of Biomolecular Complexes

Edited by

W. Ens, K. G. Standing and I. V. Chernushevich

NATO ASI Series

Series C: Mathematical and Physical Sciences – Vol. 510

New Methods for the Study of Biomolecular Complexes

edited by

W. Ens

K. G. Standing

The University of Manitoba,
Winnipeg, Manitoba, Canada

and

I. V. Chernushevich

PE SCIEX,
Concord, Ontario, Canada

Kluwer Academic Publishers

Dordrecht / Boston / London

Published in cooperation with NATO Scientific Affairs Division

Proceedings of the NATO Advanced Research Workshop on
New Methods for the Study of Molecular Aggregates
The Lodge at Kananaskis Village, Alberta, Canada
16–20 June 1996

A C.I.P. Catalogue record for this book is available from the Library of Congress.

ISBN 0-7923-5003-0

Published by Kluwer Academic Publishers,
P.O. Box 17, 3300 AA Dordrecht, The Netherlands.

Sold and distributed in the U.S.A. and Canada
by Kluwer Academic Publishers,
101 Philip Drive, Norwell, MA 02061, U.S.A.

In all other countries, sold and distributed
by Kluwer Academic Publishers,
P.O. Box 322, 3300 AH Dordrecht, The Netherlands.

Printed on acid-free paper

Printed in the Netherlands

New Methods for the Study of Biomolecular Complexes

NATO ASI Series

Advanced Science Institutes Series

A Series presenting the results of activities sponsored by the NATO Science Committee, which aims at the dissemination of advanced scientific and technological knowledge, with a view to strengthening links between scientific communities.

The Series is published by an international board of publishers in conjunction with the NATO Scientific Affairs Division

A Life Sciences	Plenum Publishing Corporation
B Physics	London and New York
C Mathematical and Physical Sciences	Kluwer Academic Publishers
D Behavioural and Social Sciences	Dordrecht, Boston and London
E Applied Sciences	
F Computer and Systems Sciences	Springer-Verlag
G Ecological Sciences	Berlin, Heidelberg, New York, London,
H Cell Biology	Paris and Tokyo
I Global Environmental Change	

PARTNERSHIP SUB-SERIES

1. Disarmament Technologies	Kluwer Academic Publishers
2. Environment	Springer-Verlag / Kluwer Academic Publishers
3. High Technology	Kluwer Academic Publishers
4. Science and Technology Policy	Kluwer Academic Publishers
5. Computer Networking	Kluwer Academic Publishers

The Partnership Sub-Series incorporates activities undertaken in collaboration with NATO's Cooperation Partners, the countries of the CIS and Central and Eastern Europe, in Priority Areas of concern to those countries.

NATO-PCO-DATA BASE

The electronic index to the NATO ASI Series provides full bibliographical references (with keywords and/or abstracts) to more than 50000 contributions from international scientists published in all sections of the NATO ASI Series.
Access to the NATO-PCO-DATA BASE is possible in two ways:

– via online FILE 128 (NATO-PCO-DATA BASE) hosted by ESRIN,
Via Galileo Galilei, I-00044 Frascati, Italy.

– via CD-ROM "NATO-PCO-DATA BASE" with user-friendly retrieval software in English, French and German (© WTV GmbH and DATAWARE Technologies Inc. 1989).

The CD-ROM can be ordered through any member of the Board of Publishers or through NATO-PCO, Overijse, Belgium.

Series C: Mathematical and Physical Sciences – Vol. 510

This book contains the proceedings of a NATO Advanced Research Workshop held within the programme of activities of the NATO Special Programme on Supramolecular Chemistry as part of the activities of the NATO Science Committee.

Other books previously published as a result of the activities of the Special Programme are:

WIPFF, G. (Ed.), *Computational Approaches in Supramolecular Chemistry.* (ASIC 426) 1994. ISBN 0-7923-2767-5

FLEISCHAKER, G.R., COLONNA, S. and LUISI, P.L. (Eds.), *Self-Production of Supramolecular Structures.* From Synthetic Structures to Models of Minimal Living Systems. (ASIC 446) 1994. ISBN 0-7923-3163-X

FABBRIZZI, L., POGGI, A. (Eds.), *Transition Metals in Supramolecular Chemistry.* (ASIC 448) 1994. ISBN 0-7923-3196-6

BECHER, J. and SCHAUMBURG, K. (Eds.), *Molecular Engineering for Advanced Materials.* (ASIC 456) 1995. ISBN 0-7923-3347-0

LA MAR, G.N. (Ed.), *Nuclear Magnetic Resonance of Paramagnetic Macromolecules.* (ASIC 457) 1995. ISBN 0-7923-3348-9

SIEGEL, JAY S. (Ed.), *Supramolecular Stereochemistry.* (ASIC 473) 1995. ISBN 0-7923-3702-6

WILCOX, C.S. and HAMILTON A.D. (Eds.), *Molecular Design and Bioorganic Catalysis.* (ASIC 478) 1996. ISBN 0-7923-4024-8

MEUNIER, B. (Ed.), *DNA and RNA Cleavers and Chemotherapy of Cancer and Viral Diseases.* (ASIC 479) 1996. ISBN 0-7923-4025-6

KAHN, O. (Ed.), *Magnetism: A Supramolecular Function.* (ASIC 484) 1996. ISBN 0-7923-4153-8

ECHEGOYEN, L., KAIFER ANGEL E. (Eds.), *Physical Supramolecular Chemistry.* (ASIC 485) 1996. ISBN 0-7923-4181-3

DESVERGNE J.P., CZARNIK A.W. (Eds.), *Chemosensors of Ion and Molecule Recognition.* (ASIC 492) 1997. ISBN 0-7923-4555-X

MICHL J., *Modular Chemistry.* (ASIC 499) ISBN 0-7923-4730-7

Preface

A NATO Advanced Research Workshop entitled New Methods for the Study of Molecular Aggregates was held at The Lodge at Kananaskis Village, Alberta, Canada from 16 - 20 June 1996. In fact the meeting was entirely concerned with the problem of analyzing biomolecular complexes, so the title of these proceedings has been altered to give a more precise description of the content. The workshop was hosted by the time-of-flight group of the Department of Physics at the University of Manitoba, and was attended by 64 participants from around the world. Twenty-one invited talks were given and 27 papers were presented as posters. Of the 48 contributions, 22 papers (12 orals, 10 posters) are included in these proceedings.

The subject of the conference was the investigation of noncovalent biomolecular complexes, with particular focus on the application of mass spectrometry to their characterization. Two new ionization techniques introduced in the late 1980s, electrospray ionization (ESI) and matrix-assisted laser desorption/ionization (MALDI), resulted in a breakthrough in mass spectrometry, enabling its use in molecular weight and primary structure determination of biopolymers larger than 100 kDa. Recently it has been discovered that ESI mass spectrometry may also be used to characterize complexes containing noncovalent interactions, thus opening new perspectives for supramolecular chemistry. ESI mass spectrometry has the advantage that the sample is introduced from a homogenous solution which can be maintained at near physiological conditions of pH, concentration, and temperature. Analysis of noncovalent complexes by MALDI mass spectrometry is more difficult because the sample is normally dissolved in a solution with an acidic matrix and the ions are desorbed from the solid phase. Both of these steps can potentially cause complexes to denature. Still, the simplicity of the MALDI spectra and its greater tolerance to impurities has motivated considerable effort to solve the problems and some progress has been made.

The first section of these proceedings contains papers which indicate the role mass spectrometry can play in combination with chemical methods in the investigation of biomolecular complexes. The next two sections focus more on the specifics of ESI and MALDI mass spectrometry as new methods to characterize these interesting entities. The final two sections describe other techniques including some emerging methods as well as the established methods like x-ray crystallography and NMR.

The Lodge at Kananaskis provided a relaxed, informal environment suitable for lively and productive discussions during the sessions as well as between them. We are grateful to NATO for providing generous financial support for the invited speakers and for a number of young researchers.

Werner Ens
Ken Standing
Winnipeg, Nov. 1997

TABLE OF CONTENTS

Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry

X-Ray Crystallography and Mass Spectrometry

Other Techniques

A GENERAL APPROACH TO DECIPHERING HOW PROTEINS WORK USING SYNTHETIC CHEMISTRY AND MASS SPECTROMETRY

MICHAEL C. FITZGERALD and STEPHEN B. H. KENT
The Scripps Research Institute
10666 N. Torrey Pines Rd.
La Jolla, CA 92037

1. Introduction

Understanding the molecular basis of protein function is currently a major focus of biological research. As chemists and biologists we would like to understand, for example, how protein domains recognize their target ligands or how enzymes are capable of catalyzing chemical reactions at nearly diffusion limited rates which in some cases are 10^6-10^{12} times faster than that of the uncatalyzed reaction. One way to get answers to questions like these is to use chemical synthesis to introduce specific mutations into the protein molecule and determine their effect on function [1].

Chemical synthesis is an especially powerful approach for studying protein structure and function. The preparation of proteins by total chemical synthesis permits the construction of protein analogues with a wide variety of both natural and unnatural modifications. Therefore by using total chemical synthesis it becomes possible to ask - and answer - very specific questions about how a protein works.

Here we describe how chemical synthesis can be used, in conjunction with mass spectrometry, to study the molecular basis of protein function. Currently, solid phase peptide synthesis (SPPS) methods permit the routine preparation of polypeptide chains up to ~40 to 60 amino acids [2]. The total chemical synthesis of larger protein constructs (up to 200 amino acid residues in length) has also been possible using recently developed chemical ligation strategies for the chemoselective ligation of unprotected peptide segments [3]. Thus, small proteins and enzymes in the 40-200 amino acid size range are accessible by total chemical synthesis. The ability to synthesize proteins of this size is of quite general utility because the fundamental building blocks of the protein world are 'functional domains' that are typically 110 ± 40 amino acids. As an example of how total chemical synthesis can be used to study the structure and function of an enzyme, we highlight our work on 4-oxalocrotonate tautomerase (4OT).

W. Ens et al. (eds.), New Methods for the Study of Biomolecular Complexes, 1–16.

2. Probing the Function of an Enzyme by Total Chemical Synthesis.

4OT is an enzyme used by certain soil bacteria to help degrade aromatic hydrocarbons to provide intermediates for the Krebs cycle [4,5]. The 4OT enzyme, which contains no cysteine residues, is a hexamer of identical 62 amino acid polypeptide chains and catalyzes the 1,3-allylic isomerization of 2-oxo-4-*(E)*-hexenedioate (**1**) to 2-oxo-3-*(E)*-hexenedioate (**3**) through the intermediate 2-hydroxymuconate (**2**) (Figure 1A). Recent X-ray crystallographic data on a 4OT isozyme has shown that the homo-hexameric, enzyme complex is a trimer of dimers; each dimer contains a 4 stranded beta sheet region and 2 anti-parallel helices [6]. Results from inhibition studies with an active site-directed irreversible inhibitor of 4OT have suggested that the enzyme complex contains six active sites and have implicated the N-terminal proline as the catalytic base [7]. It has also been hypothesized from the NMR solution structure and the three dimensional crystal structure of 4OT that two Arg residues (Arg^{11} of one subunit and Arg^{39} of an adjacent subunit) are important for binding the dicarboxylic acid substrate (**1**) in the active site of the enzyme (Figure 1B) [6,8].

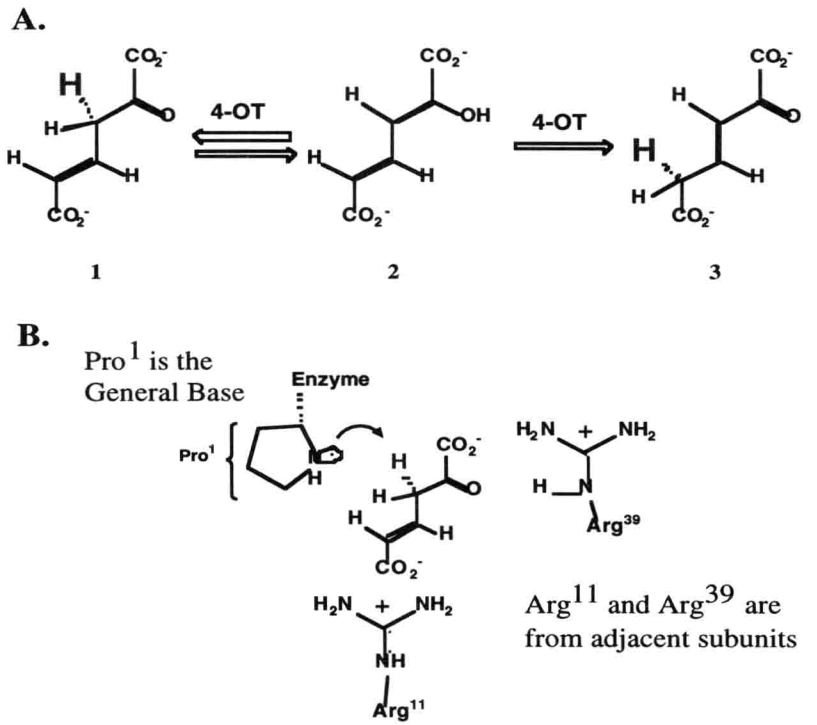

Figure 1. 4OT-catalyzed isomerization reaction (A) and proposed mechanism of action (B).

The X-ray crystallographic results and biochemical studies on 4OT that are referenced above have each helped identify catalytically important residues in the enzyme; however, the molecular basis for the function of these residues is not well understood. Currently, we are using total chemical synthesis to prepare 4OT analogues with various chemical modifications designed to address specific questions about the catalytic mechanism of this highly efficient enzyme.

We have initially focused on the synthesis and characterization of four different 4OT analogues including: (*des*Pro^1)4OT, a truncated construct in which Pro^1 was deleted; (Cpc^1)4OT in which Pro^1 was replaced with cyclopentane carboxylate; a derivative [Met(O)^{45}]4OT in which the side chain sulfur of Met^{45} was oxidized to the sulfoxide; and an analogue (Nle^{45})4OT in which Met^{45} was replaced with norleucine. Our purpose in studying (Met(O)^{45}]4OT was to determine what effects the oxidation of the single Met residue in 4OT had on the enzyme's structure, because enzyme preparations containing partially oxidized material displayed reduced activity. The (Nle^{45})4OT construct was designed to eliminate the possibility of oxidizing the enzyme during sample handling. The (*des*Pro^1)4OT and (Cpc^1)4OT constructs were designed to investigate the catalytic significance of the N-terminal proline reside. It is noteworthy that the (Cpc^1)4OT analogue is only accessible by total chemical synthesis methods.

2.1. SYNTHESIS AND CHARACTERIZATION OF 4OT ANALOGUES

2.1.1. *Chemical Synthesis, Purification, and Folding of Proteins*

The 62 amino acid polypeptide chain of wild-type 4OT was synthesized from protected amino acids in stepwise fashion by highly optimized, solid phase peptide synthesis (SPPS) methods using *in situ* neutralization protocols for t-butoxycarbonyl (Boc) chemistry [9]. The crude polypeptide product was purified by reversed-phase HPLC. Typically, 50-100 mg of high purity 62mer could be obtained from a single 0.2 mmol synthesis (the smallest convenient scale of laboratory synthesis). Analytical reversed-phase HPLC and electrospray ionization mass spectrometry (ESI-MS) were used to confirm the identity and purity of the final polypeptide product. Results for the HPLC and ESI-MS analysis of purified, synthetic 4OT monomer are shown below in Figure 2. The purified synthetic peptide product eluted as a single peak using a shallow gradient (40-60% acetonitrile/water in 0.1% triflouroacetic acid over 30 minutes). The monomeric molecular weight of the 4OT sample was 6809.9 $\pm$ 0.7 (as determined by ESI-MS) and in good agreement with the expected mass, 6810.7 (average isotopic composition). The monomer polypeptide chains of 4OT were folded to form the homo-hexameric enzyme complex during a 2 hour, room temperature incubation in assay buffer (20 mM sodium phosphate, pH 7.4).

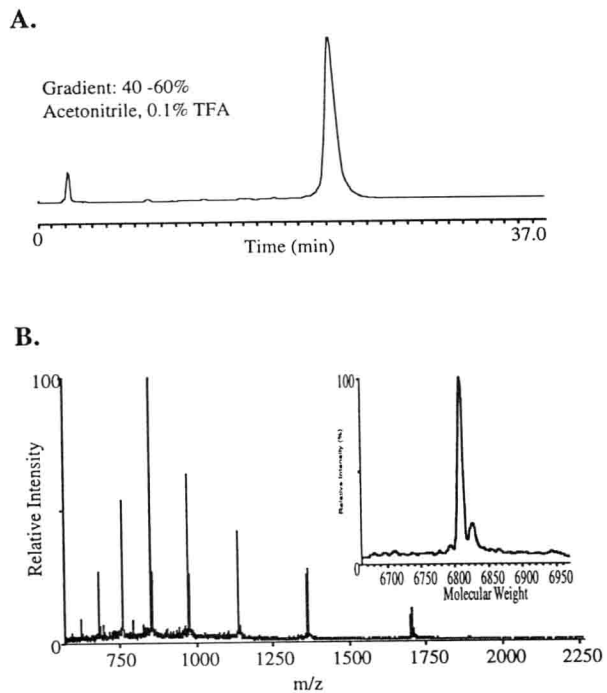

Figure 2. (A) Analytical reversed-phase HPLC analysis of purified 62 amino acid monomer of 4OT. (B) Electrospray mass spectra of HPLC purified 4OT.

In addition to the wild-type enzyme, we have also used total chemical synthesis to prepare the (*des*Pro^1)4OT, (Cpc^1)4OT, [Met(O)^{45}]4OT, and (Nle^{45})4OT analogues. The polypeptide chains of the (*des*Pro^1)4OT, (Nle^{45})4OT, and (Cpc^1)4OT constructs were each synthesized in stepwise fashion using SPPS methods. In the case of (Cpc^1)4OT, unprotected cyclopentanecarboxylic acid (Cpc) was used instead of Boc-proline in the last coupling reaction of the synthesis. The [Met(O)^{45}]4OT polypeptide chain was prepared by mild oxidation (0.3% v/v hydrogen peroxide in 0.1% triflouroacetic acid) of purified, synthetic 4OT. Analysis of the reaction by ESI-MS revealed that oxidation to the sulfoxide was complete and no sulfone was observed. Each 4OT analogue was folded in assay buffer as described above for the wild-type enzyme.

2.1.2. *Circular Dichroism Spectroscopy*

Circular dichroism (CD) spectroscopy was used to characterize the secondary structure

of each folded 4OT analogue. The CD spectra recorded for the wild-type enzyme and for each analogue are shown in Figures 3A and B. The helical content of each analogue was estimated from the magnitude of its molar ellipticity at 222 nm [10]. The percent helicities calculated for 4OT, (Nle^{45})4OT, (Cpc^{1})4OT and [Met(O)^{45}]4OT were each

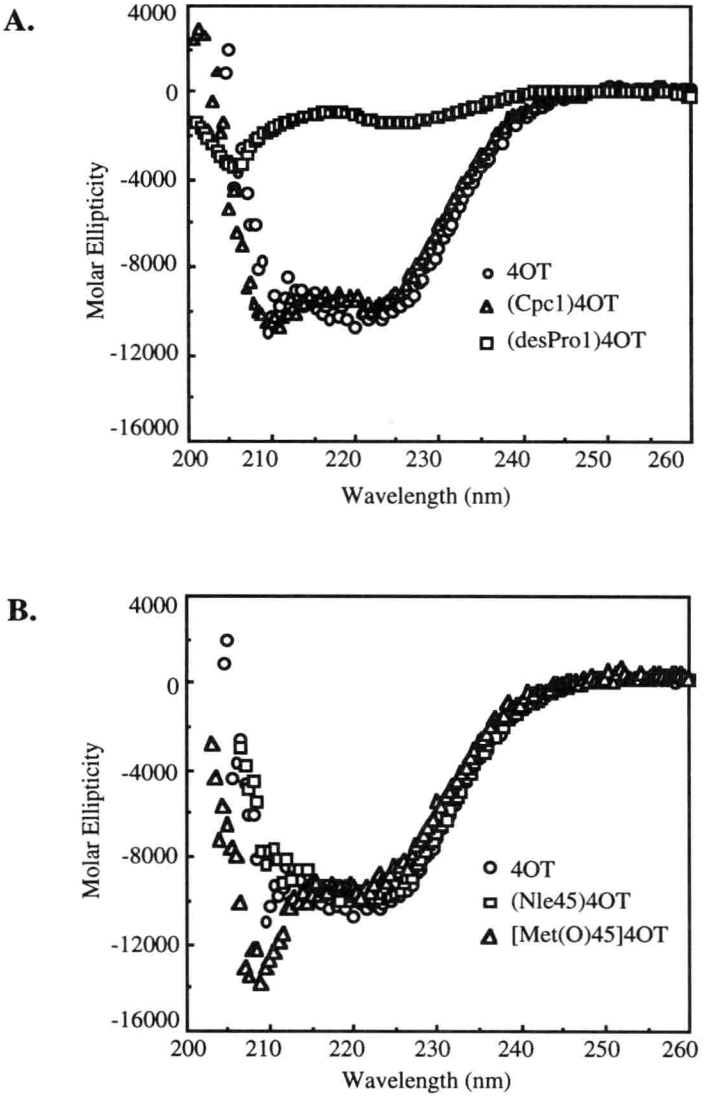

Figure 3. CD Spectra for 4OT and Analogues

approximately 26% and in relatively good agreement with that previously reported for recombinant 4OT (21%) [11]. Furthermore, the other features of the CD spectra recorded for (Cpc^{1})4OT (see Figure 3A) and (Nle^{45})4OT (see Figure 3B) were essentially identical to the spectrum recorded for 4OT indicating that the secondary structures of these analogues are similar to those of the wild-type enzyme. In contrast, the unique minimum at 208 nm observed in the CD spectrum of [Met(O)^{45}]4OT (see Figure 3B) suggests that the secondary structure of this analogue differs in some aspects from that of the wild type enzyme. The essentially featureless spectrum of the (*des*Pro^{1})4OT analogue (see Figure 3A) showed that this polypeptide has very little defined secondary structure.

2.1.3. *Electrospray Ionization Time-of-Flight (ESI-TOF) Mass Spectrometry*

ESI-TOF mass spectrometry was used to assess the oligomeric state of each 4OT analogue. The noncovalent hexamer of 4OT is readily detected by ESI-TOF at m/z values between 3000 and 3500, and the technique has proven very sensitive for examining the effects that subtle differences in covalent structure can have on an enzyme's higher order structure [12]. The ESI-TOF mass spectra recorded for the 4OT analogues in this study are shown in Figure 4 (a typical spectrum of wild-type 4OT is also included for comparison).

The spectra in Figure 4 were acquired under identical "native" electrospray conditions using the same buffer (5 mM ammonium bicarbonate, pH 7.5), needle position, and declustering potential. Strong hexamer signals were detected for the (Nle^{45})4OT and (Cpc^{1})4OT samples indicating that 62 residue polypeptide chains of these analogues each fold in solution to form a hexameric enzyme complex. The absence of any multimeric species in the native mass spectrum of (*des*Pro^{1})4OT (Figure 4E) suggests that this truncated analogue has no discernible quaternary structure. The presence (albeit weak) of multiply charged hexamer ions in the native mass spectrum of [Met(O)^{45}]4OT (Figure 4D) indicates that this construct is capable of forming a homo-hexameric complex. However, the ratio of hexamer to monomer in the [Met(O)^{45}]4OT spectrum was significantly smaller (~50-fold less) than the hexamer to monomer ratio in the 4OT, (Nle^{45})4OT, and (Cpc^{1})4OT spectra. These results suggest that the noncovalent interactions defining the hexameric [Met(O)^{45}]4OT complex are significantly weaker than those in 4OT.

2.1.4. *Enzymatic Activity*

The catalytic efficiency of each 4OT analogue in this study was determined by studying the kinetics of the conversion of 2-hydroxymuconate (**2**) to 2-oxo-3-trans-hexenedioate (**3**) (see Figure 1). The K_M, and k_{cat} values determined for the wild-type enzyme and each analogue are given in Table 1.

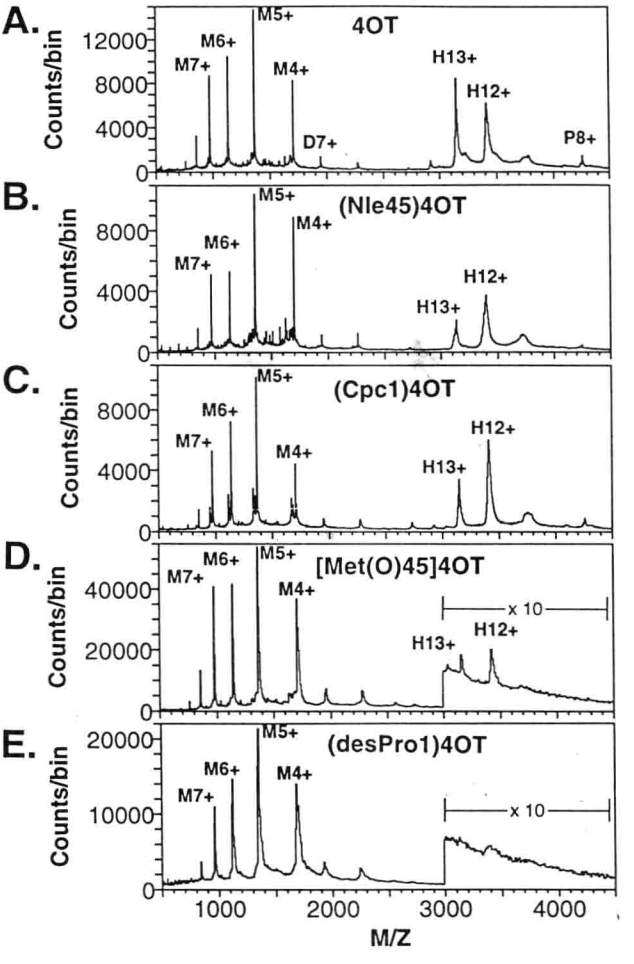

Figure 4. ESI-TOF Analysis of 4OT Analogues

The kinetic data show that the (Nle^{45})4OT and [Met(O)^{45}]4OT analogues displayed enzymatic activity comparable to that of wild-type 4OT (k_{cat} ~2900 s^{-1}). By contrast, for the (Cpc^{1})4OT and (*des*Pro^{1})4OT analogues no rate enhancement over the nonenzymatic chemical ketonization (uncatalyzed rate: 8.7 x10^{-4} s^{-1}) of 2-hydroxymuconate to 2-oxo-3-trans-hexenedioate was detected (detection limit: kcat ~1 s^{-1}). It is also noteworthy that the full enzymatic activity of [Met(O)^{45}]4OT reported in Table 1 was only obtained when enzyme stock solutions at concentrations greater than 40 μM (based on total amount of monomer). The [Met(O)^{45}]4OT analogue is