Doctoral Dissertation

Exploration of Biomolecularly Transparent IR Region for Structural Identification of Biomolecules Using VCD

（生体分子の構造同定を目指した生体分子非吸収スペクトル領域におけるVCD研究）

Mohamad Zarif Bin Mohd Zubir

Graduate School of Life Science, Hokkaido University

Transdisciplinary Life Science Course

Laboratory of Molecular Chemical Biology

2022,SEPTEMBER

TABLE OF CONTENTS

LIST OF SYMBOLS AND ABBREVIATIONS

—	:	Single bond
°	:	Degree
°C	:	Degree Celsius
=	:	Double bond
≡	:	Triple bond
^{12}C	:	Carbon-12
^{13}C	:	Carbon-13
^{1}H	:	Proton
2D	:	2 dimension
6LG	:	6-*O*-lauroylated D-glucose
A	:	Absorbance
α	:	Alpha
Å	:	Angstrom
Ac	:	Acetate
Ac_2O	:	Acetic anhydride
AcCl	:	Acetyl chloride
AcOH	:	Acetic acid
aq	:	Aqueous
Ara	:	Arabinose
ArH	:	Aryl
β	:	Beta
BaF_2	:	Barium fluoride
Bz_4G	:	Tetrabenzoly D-glucose
BzCl	:	Benzoly chloride

c	:	Concentration
C	:	Carbon
C-O	:	Ether
C=N	:	Imines
C=O	:	Carbonyl
C=S	:	Thiocarbonyl
C≡CH	:	Alkyne
C≡N	:	Cyano
ca.	:	circa
CaF_2	:	Calcium fluoride
CCl_4	:	Carbon tetrachloride
CD	:	Circular dichroism
CD_3	:	Deuterated methyl
CD_3OD	:	Deuterated methanol
$CDCl_3$	:	Deuterated chloroform
CH_2Cl_2	:	Dichloromethane
$CHCl_3$	:	Chloroform
cm	:	Centimetre
CO_2	:	Carbon dioxide
cryoEM	:	Cryo-electron microscopy
D	:	Dipole strength
δ	:	Chemical shift values
d	:	Doublet
D	:	Dextro
D_2O	:	Deuterated water

Da	:	Dalton
DCM	:	Dichloromethane
DFT	:	Density functional theory
DFT	:	Density functional theory
DMSO-d_6	:	Deuterated dimethyl sulfoxide
DNA	:	Deoxyribonucleic acid
ECD	:	Electronic circular dichroism
EDTMs	:	Electric dipole transition moments
ESR	:	Electron spin resonance
Et_2O	:	Diethyl ether
Et_3N	:	Triethylamine
EtOAc	:	Ethyl acetate
ϕ	:	Diameter
f	:	Furanose
FRET	:	Fluorescence resonance energy transfer
FT	:	Fourier transform
Fuc	:	Fucose
Gal	:	Galactose
Glc	:	Glucose
h	:	Hour
H	:	Hydrogen
H_2O	:	Water
H_2O_2	:	Hydrogen peroxide
HCl	:	Hydrochloric acid
HRMS	:	High resolution mass spectrometry

InSb	:	Indium antimonide
IR	:	Infrared
J	:	*J* coupling
K_2CO_3	:	Potassium carbonate
Kcal	:	Kilocalorie
kJ	:	KiloJoules
l	:	Path length
L	:	Levo
m	:	Magnetic transition moment
M	:	Molar
m	:	Multiplet
M.W.	:	Molecular weight
m/z	:	Mass-to-charge ratio
Man	:	Mannose
MCT	:	Medium-chain triglycerides
MD	:	Molecular Dynamics
MeCN	:	Acetonitrile
MeOH	:	Methanol,
MeONa	:	Sodium methoxide
Methyl-d_3	:	Deuterated methyl
mg	:	Milligram
mg	:	Milligram
$MgSO_4$	:	Magnesium sulfate
MHz	:	Megahertz
microED	:	Micro-crystal electron diffraction

mL	:	Millilitre
mM	:	Millimolar
MM	:	Molecular mechanics
MMFF	:	Merck molecular force field
Mmol	:	Millimole
MS	:	Mass spectrometry
N	:	Nitrogen
ν	:	Frequency
N=N=N	:	Azido
N≡C	:	Isocyano
N_2	:	Nitrogen gas
$Na_2S_2O_3$	:	Sodium thiosulfate
NaH	:	Sodium hydride
$NaHCO_3$	:	Sodium bicarbonate
NaN_3	:	Sodium azide
$NaNO_2$	:	Sodium nitrite
NaOH	:	Sodium hydroxide
NH_3	:	Ammonia
NH_4Cl	:	Ammonium chloride
nm	:	Nanometre
NMR	:	Nuclear magnetic resonance
OCD_3	:	Deuterated methoxy
π	:	Pi
p-TsOH•H_2O	:	p-Toluenesulfonic acid
PCM	:	Polarizable continuum model

Pd	:	Palladium
PEM	:	Photoelastic modulator
θ	:	Dihedral angle
QM	:	Quantum mechanics
R	:	Rotational strength
r	:	Distance
R	:	Substituent group
R	:	Rectus
RNA	:	Ribonucleic acid
ROA	:	Raman optical activity
rt	:	Room temperature
S	:	Sinister
s	:	Singlet
SiO_2	:	Silicon dioxide
$SOCl_2$	:	Thionyl chloride
SPR	:	Surface plasma resonance
TBAF	:	Tetra-n-butylammonium fluoride
TBSCl	:	tert-Butyldimethylchlorosilane
Tf_2O	:	Trifluoromethanesulfonic anhydride
THF	:	Tetrahydrofuran
VCD	:	Vibrational circular dichroism
VIS-UV	:	Visible-ultraviolet
VROA	:	Vibrational Raman optical activity
VUV-CD	:	Vacuum-ultraviolet circular dichroism
Xyl	:	Xylose

ε	:	Molar absorptivity
μ	:	Electric dipole transition moment
µL	:	Microlitre
µm	:	Micrometre
µmol	:	Micromole

CHAPTER 1: INTRODUCTION

1.1 General background: Structural and functional relationship of biomolecules.

Some of the most abundant classes of biomolecules present in the biological system are carbohydrates, lipids, nucleic acids, and proteins. They act as primary building blocks in performing all the essential functions in biological processes.[1] The diversity of biomolecules in terms of their structural formations which can be rather complex and complicated to identify has been a great challenge for researchers worldwide. The structural characteristic of biomolecule is crucial to be fully understood to completely identify their vital functionality in an organism.

As an example, the double helical structure of deoxyribonucleic acid (DNA) was discovered by James Watson and Francis Crick in 1953 from the X-ray diffraction pictures taken by Rosalind Franklin [2]. DNA is presents in the nucleus of the eukaryotic organisms and is a complex supercoiled biomolecule that served two fundamental roles. First, they control the maintain the genetic information and passing this information to each new cell. This is made possible through the presence of the double helix structural arrangement, where their base paring nucleotides are arranges in specific and selective manner complementing each exclusive function. Second, the genetic code present in the DNA also can undergo transcription into RNA (ribonucleic acid) which can then be translated into various specific proteins. Their infinite arrays of sequencing pattern and structurally flexible conformation that are polymorphic in nature enable this complex helical structure to perform all these bioprocesses. [3]

DNA display a central example representing the relationship between structural conformation of biomolecules and its functions. However, more complex biological

phenomena were discovered in daily basis which cannot be fully understood until proper characterization of the structure of reacting biomolecules is comprehensively investigated. Therefore, the development of techniques and methods used for characterization and identification of biomolecules is an urgent necessity.

1.2 Importance analytical instrumentation techniques for structural identification of biomolecules

The ever advancing and rapid development in the field of analytical instrumentation techniques permits the ability of effective with high precision discovery of complex biological molecules at miniscule level. Nevertheless, even with availability of the existing analytical instrumentation, there are still increasing demand for new analytical technologies with constant improvements in terms of the sensitivity, accuracy, and specificity. Within the near future, it is expected for rising of more new development of, innovative analytical technologies and methodologies. It is particularly needed for the research progression in the field of biochemistry and pharmaceutical for early drug development stage. These tools should be able to extensively employ detection of vast array of biological molecules such as metalloenzymes, oligonucleotides, glycolipids, and so on.

Currently, complex biomolecules sequence and composition has been mostly elucidated through several methods including NMR and MS techniques. NMR has been use for study of large complex biomolecules including the oligomerization the β-Amyloid peptide, glycoprotein and supramolecular phospholipids. [4,5] However, the measurement and interpretation of NMR spectroscopy often requires a long time and demanded specialized expertise to operate the instrument to be practically applied on complicated

biomolecules. Nowadays, the complete characterization of biomolecules requires the combination of multiple sophisticated analytical techniques which includes surface plasma resonance (SPR),[6] X-ray crystallography,[7] electronic circular dichroism (ECD) spectroscopy,[8] fluorescence resonance energy transfer (FRET),[9] and Raman optical activity (ROA)[10] are just some to mention. Instrumentation techniques can also be further backed by computational calculation methods that utilized an in-depth knowledge at molecular level which can pave the essential understanding of bioprocesses and developments of new drugs or medicine for the betterment of mankind.

Comparing to synthetic molecule, biomolecules are made up of structures that can be less straightforward due to their complicated nature. Hence, the data analysis of uncommon biomolecules has been rather difficult as the method depends heavily on the comparison of previously observed fragmentation patterns from a library of database. Due to these limitations, high demand existed for further development of new analytical tools. Addition to that, most current analyses are based on intrinsically achiral theories. This approach limits the characterization of biomolecules which is commonly made up of an assembly of chiral centres. Further development of analytical tool to effectively study chiral biomolecules such as carbohydrates and peptides scaffold for instant, will be advantageous for extracting their stereochemical information. In this study, we expanded and explored the potential of VCD as an instrument for the characterization of biomolecules.

1.3 Vibrational circular dichroism (VCD)

Vibrational circular dichroism (VCD) is described as circular dichroism (CD) in vibrational transitions in the IR region of chiral molecules. VCD operation is based on the measurement of optical rotation or optical rotatory dispersion function of wavelength that can be either clockwise or counterclockwise, of the optical activity possessed by the chiral samples. These chiroptical techniques are most widely used in biomolecular study because of their reliability to determine absolute configuration of a molecule with the aid of theoretical calculation.

Meanwhile, circular dichroism (CD) measures the absorption differential of left versus right circularly polarized radiation. Traditionally, "CD" has been used for the circular dichroic measurement in the visible-ultraviolet (VIS-UV) wavelengths and the CD phenomenon commonly exists throughout all spectral regions (**Figure 1.1**). However, conventional "CD" (i.e., VIS-UV CD) or also known as electronic circular dichroism (ECD) since the absorption originates from electronic transitions, is distinguishable from recently developed vibrational circular dichroism (VCD), which targets the infrared (IR) region. On the other hand, vacuum-ultraviolet circular dichroism (VUV-CD) is a class of CD measurement conducted below 190 nm. Vibrational Raman optical activity (VROA) can be regarded as a vibrational Raman counterpart, which has been recently evolved parallel with VCD.[11]

Absorption in the corresponding spectral region is necessary to analyze biomolecules by using CD. Thus, not all chiral biomolecules can be easily characterized by just relying on CD alone. In addition, the difficulties in data interpretation due to broadened spectral features of CD for biomolecule structural analysis remains to be resolved. Due to the absence of VIS-UV absorption for some biomolecules, manipulation related to chromophores introduction is necessary for ECD study. Recently, the

emergence of hi-tech VCD instrumentation for chiroptical analysis that allow the observation of detailed structural information through a wider frequency region without the need of any derivatization since all organic molecules absorb IR radiation. Variety areas from small chiral molecules to biomacromolecules widely used VCD as analysis tools.[11]

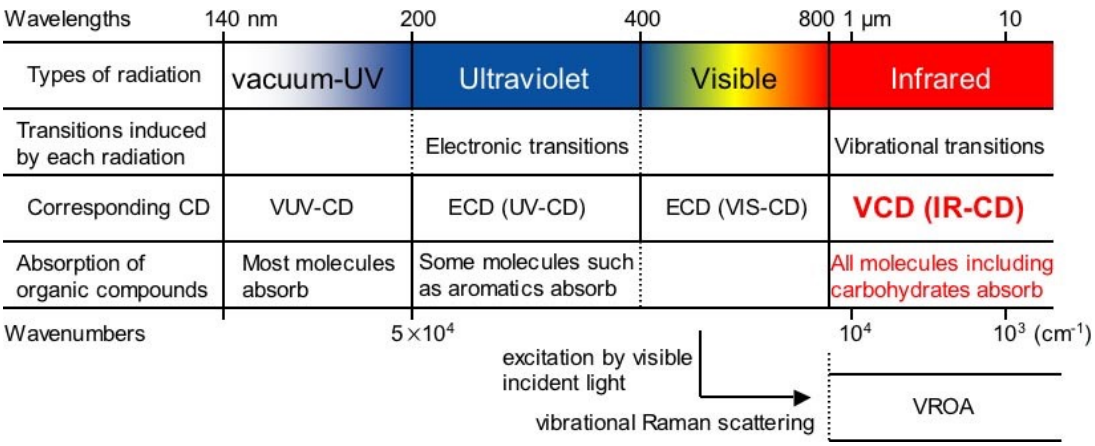


Figure 1.1: Types of radiation, the induced transitions, and the corresponding CD. VCD is considered as an extended version of conventional ECD. While VUV-CD (vacuum-ultraviolet circular dichroism) is conducted below 200 nm, and VROA (vibrational Raman optical activity) can be regarded as a vibrational Raman counterpart. The advantage of VCD comes from spectroscopic detail of IR with the stereochemical sensitivity of CD.[11]

1.4 The theoretical aspect of VCD

VCD operates by measuring the difference in the absorbance of a chiral sample for left (A_L) minus right (A_R) circularly polarized IR incident light.

$$\Delta A = A_L - A_R$$

where A is the decadic absorbance given by the function of the average transmission with and without the sample, I and I_0, respectively.

$$A = -\log_{10}(I/I_0)$$

Ordinary IR absorption is expressed as the average of A_L and A_R.

$$A = (A_L + A_R)/2$$

A (+)-enantiomer and its (-)-antipode exhibit VCD of opposite sign with equal intensity due to all the optical activity phenomenon and responded equivalently to unpolarized light, namely,

$$\Delta A(+) = -\Delta A(-)$$

$$A(+) = A(-).$$

The dipole strength D is proportional to the intensity of an IR absorption band, which is the absolute square of the electric dipole transition moment, μ, of the molecule for each normal mode of vibration.

$$D = |\mu|^2$$

As an example, the strong carbonyl IR absorption is the result of the large difference in electronegativity between carbon and oxygen that causes in high magnitude of electric dipole transition moment of the C=O stretching. Furthermore, the rotational strength, R is proportional to VCD intensity, an imaginary part of the scalar product of the electric and magnetic transition moments, respectively, μ and m.

$$R = \mathrm{Im}(\mu \bullet m)$$

Prediction of the intensity and sign of a VCD absorption band is not possible without theoretical simulation, the magnitude and direction of these transition moments. Applying the Beer-Lambert's law, VCD is generally presented in terms of the difference in molar absorptivity ε,

$$\Delta\varepsilon = \varepsilon L - \varepsilon R = \Delta A/cl$$

The molar circular dichroism Δε is a more useful quantity than raw VCD intensity ΔA for the comparison of experimental measurements with theoretical calculations. In this expression the effects of concentration c and path length l may be cancelled. For biomolecular samples, the Δε representation is often impractical not only because the accurate concentration is hard to know but also because they sometimes consist of more than one component. The ratio of the VCD intensity to the IR intensity, called the anisotropy ratio, g, is consistent in both the ΔA and Δε proportions.

$$g = \Delta A/A = \Delta\varepsilon/\varepsilon = 4R/D$$

The g-factor can also be obtained as four times the rotatory strength R divided by the dipole strength D.[11]

VCD and IR spectra can also be simulated computational calculation base on the quantum mechanics for the predictable energy vibrational mode of each atom. Theoretical VCD calculation algorithm with high accuracy is mostly available in commercial and free software have become a convenient tool for all researchers. The theoretical calculation is carried out independently of the experimental measurement and to be compared with it. The comparison of calculated and observed spectra of a target chiral molecule informs its absolute configuration, conformation.[12–17] VCD spectroscopy is now one of the most convenient and reliable tools for nonempirical determination of the absolute configuration of natural and synthetic chiral compounds, due to the convenience and high reliability of the theoretical prediction.[11]

Theoretical calculation starts from a conformation search at the molecular mechanics level. The stable conformers are further optimized to obtain the minimum energy geometry at the density functional theory (DFT) level. The B3LYP or B3PW91 functionals and 6-31G(d) basis set are being widely used. For the optimized low-energy conformations, the VCD and IR spectra are calculated at the same level of theory. If the molecule adopts more than one conformer, the VCD and IR spectra for each stable conformer must be calculated and superposed by the Boltzmann weighted population.[11] The absolute stereochemistry of biomolecule can be determined through this procedure.

1.5 The experimental aspect of VCD

All current commercial VCD spectrometers are targeting the mid-IR region, and only a part of them has accessibility to the hydrogen stretching region. In the past decade, VCD instruments adopt the Fourier transform (FT) optical system as illustrated in **Figure 1.3.** In this system, the radiation beam from a broadband thermal IR source first directs through a Michelson interferometer, as an ordinary FT-IR spectrometer. The exiting IR beam is polarized linearly and then modulated by a photoelastic modulator (PEM) to left- and right-circularly polarized state in the frequency range of 60-30 kHz. The circularly polarized light passes through the chiral sample and then focused with a lens onto a liquid-nitrogen cooled detector. The electronic signal at the detector is processed to yield raw VCD spectra by subsequent electronics. In this process, the corresponding IR spectra can be obtained simultaneously.[11]

Small intensities of VCD absorption have been a major disadvantage. Even with a state-of-the-art instrument, the measurement requires a longer acquisition time and a higher sample amount (typically 1-3 hours and 1-10 mg, respectively), limiting the application to rare biomolecules. A milligram sample is required to adjust the decadic absorbance (A) of a band of interest between 0.1 and 1.0, a generally recommended value for VCD measurements. Another difficulty of biomolecules is the lack of its enantiomer. VCD spectra of a chiral molecule could not be obtained without measuring the opposite enantiomer or racemic mixture because of artifacts associated with absorption bands. Such artifacts must be removed or, at least, minimized by optimizing the instrumental condition. If both enantiomers are available, even in case they do not show mirror symmetrical VCD data, the most reliable VCD spectra can be obtained by enantiomer correction.

$$\Delta A(+)_{corrected} = 1/2\ [\Delta A(+)_{raw} - \Delta A(-)_{raw}]$$

or, correction with a raw VCD spectrum of racemate.

$$\Delta A(+)_{corrected} = \Delta A(+)_{raw} - \Delta A(\pm)_{raw}$$

This approach is not applicable for biological molecules whose enantiomer is usually unavailable. An alternative and more general way to correct a solution-state sample spectrum is solvent correction.

$$\Delta A_{corrected} = \Delta A_{sample} - \Delta A_{solvent}$$

Unless the sample is not neat liquid or solid-state, an IR spectrum should be corrected with the solvent IR spectrum as well. All sugar VCD and IR spectra shown in this review are corrected by the solvent spectra.[11]

VCD samples must be isotropic since linear dichroism (birefringence) causes distortion of CD spectra, and so compounds are usually measured in solution state placed in an appropriate optical cell. The window material must also be isotropic, permeable to the wavelength range of interest, and resistant for H_2O and D_2O, which are common solvents for biomacromolecules. A BaF_2 or CaF_2 cell with a pathlength of 5-200 μm is used for routine VCD measurements. BaF_2 offers wider spectral range up to ~750 cm^{-1} but it is slightly more deliquescent. The long exposure to aqueous solution or moisture causes a clouding on the BaF_2 window surface. On the other hand, CaF_2 windows have no substantial problem with water, but preclude measurement below 1000 cm^{-1} (Fig. 4). Measurements of solid sample such as suspension in oil (31) and recent film technique (32–34) are possible but the measurement must be repeated by rotating the sample 45° or 90° around the light beam axis to verify the independence from the linear dichroic interference. The solvent should not have significant absorption in the region of measurement. For their permeability and little solvent influence, CCl_4 and $CDCl_3$ are ideal choices. The lack of hydrogen bonding properties makes these solvents favorable for calculation assisted VCD studies. Hydrophilic molecules are typically measured in water, deuterium oxide or DMSO-d_6. Biomacromolecular secondary structure analyses

via VCD have been studied in an aqueous environment. Measurements in H_2O entail difficulties arising from its strong absorption. The pathlength must be reduced to around 10 μm and, as a result, the sample must be prepared in a higher concentration, which is possible only if the solubility is high enough. A broad absorption band centered at around 1650 cm^{-1} with absorbance of ~0.5 at 6 μm pathlength is also problematic because the accompanying noise tends to obscure-1true VCD signals around 1700-1600 cm^{-1}. Alternatively, D_2O has less absorption in this region, facilitating the C=O stretch VCD measurement at a longer pathlength (25-100 μm), at a cost of spectral range below 1250 cm^{-1}. The measurements must be preceded by lyophilization twice from D_2O to avoid expected noise associated with H-D exchange. DMSO-d_6 ensures the measurements up to 1100 cm^{-1} even at a longer pathlength (200 μm is still acceptable) and has strong solubility for hydrophilic as well as amphiphilic molecules, which makes it possible to study several kinds of molecules in C-H stretching (2700-3100 cm^{-1}), C=O stretching, and lower frequency region.[11]

1.6 Research objective

In this study we aim to explore some theoretical and experimental aspects of VCD, and its application as a tool for structural analysis specifically focusing on chromophores group in the biomolecularly IR region (i.e., IR absorption in the 2300-1900 cm^{-1} region). The chromophores will be introduced as a probe for biomolecules structural characterization and identifications. First, we will focus on chromophores with X≡Y and X=Y=Z systems such as nitrile, isonitrile, alkyne, and azido groups to identify suitable groups that exhibit VCD couplets in the 2300-1900 cm^{-1} region by using a chiral binaphthyl scaffold. The incorporation of VCD exciton chirality method which makes data interpretation becomes easier by observing a strong VCD couplet originating from the interaction of two IR chromophores, whose sign reflects the absolute configuration of the molecule. We investigate the application of the bisignate VCD signals of these chromophores that is advantages toward exhibiting characteristic IR absorptions in a spectrally transparent region. We search for the best chromophore pair that is most promising for extracting structural information of biomolecules and will show the most simple and strong VCD signals whose pattern is readily predicted by computational calculations. The most suitable chromophores group will be then installed onto some biomolecules to further test their potential and applicability. Second, we utilize the C-D stretching in this region to explore the practicality of CD_3 groups for obtaining the local chirality of sugar molecules. A series of pyranosides is used as a model study. The epimeric pairs are expected to present VCD patterns in the 2300-1900 cm^{-1} region irrespective of other chiral centers. The application of this concept will be further demonstrated to other complex systems by studying the VCD spectra of the C-1 epimeric pairs of other complex sugar system such disaccharide gentiobiose, 6-lauroylated surfactant, and 5-membered glucofuranosides.

1.7 Reference

1. L. A. Urry, M. L. Cain, S. A. Wasserman, P. V. Minorsky, and R. Orr, *Campbell Biology*, 2009, **9**, 68-91
2. J. D. Watson, and F. H. Crick. *Nature*, 2003, **421**, 397-8
3. M. Kaushik, S. Kaushik, K. Roy, A. Singh, S. Mahendru, M. Kumar, S. Chaudhary, S. Ahmed, and S. Kukreti, *Biochemistry and Biophysics Reports*, 2016, **5,** 388-395
4. A. Barnes, A. J. Robertson, J. M. Louis, P. Anfinrud, and A. Bax, *J. Am. Chem. Soc*. 2019, **141**, 13762–13766
5. P. Nitschke, S. Lodge, T. Kimhofer, R. Masuda, S. -H. Bong, D. Hall, H. Schäfer, M. Spraul, N. Pompe, T. Diercks, G. Bernardo-Seisdedos, J. M. Mato, O. Millet, D. Susic, A. Henry, E. M. El-Omar, E. Holmes, J. C. Lindon, J. K. Nicholson, and J. Wist, *Anal. Chem.* 2022, **94**, 1333–1341
6. L. C. Smith, P. T. Bremer, C. S. Hwang, B. Zhou, B. Ellis, M. S. Hixon, and K. D. Janda, *J. Am. Chem. Soc.* 2019, **141**, 10489–10503
7. S. L. Heinz-Kunert, A. Pandya, V. T. Dang, P. N. Tran, S. Ghosh, D McElheny, B. D. Santarsiero, Z. Ren, and A. I. Nguyen**,** *J. Am. Chem. Soc.* 2022, **144**, 7001–7009
8. G. Pescitelli, L. D. Bari and N. Berova, *Chem. Soc. Rev.,* 2014, **43**, 5211-5233
9. S. Liu, Z. Liu, H. Wang, J. Wang, R. Liu, K. Wang and J. Huang, *Chem. Commun.,* 2022, **58**, 1414-1417
10. P. Morgante, H. D. Ludowieg, and J. Autschbach, *J. Phys. Chem. A* 2022, **126**, 2909–2927
11. T. Taniguchi, and K. Monde, *Trends in Glycoscience and Glycotechnology,* 2007, **19**, 147–164
12. T. B. Freedman, X. Cao, R. K. Dukor, and L. A. Nafie, *Chirality*, 2001, **15**, 743–758
13. T. Furo, T. Mori, T. Wada, and Y. Inoue, *J. Am. Chem. Soc.*, 2005, **127,** 8242–8243
14. K. Monde, T. Taniguchi, N. Miura, C. S. Vairappan, and M. Suzuki, *Chirality,* 2006, **18**, 335–339
15. K. Monde, N. Miura, M. Hashimoto, T. Taniguchi, and T. Inabe, *J. Am. Chem. Soc.,* 2006, **128,** 6000–6001

16. K. Monde, T. Taniguchi, N. Miura, C. S. Vairappan, and M. Suzuki, *Tetrahedron Lett.*, 2006, **47,** 4389–4392
17. H. Izumi, S. Futamura, N. Tokita, and Y. Hamada, *J. Org. Chem.,* 2007, **72**, 277–279

CHAPTER 2: EXPLORATION OF CHROMOPHORES FOR VCD COUPLET IN BIOMOLECULARLY TRANSPARENT INFRARED REGION

2.1 Abstract

Interactions of two chromophores such as carbonyl groups yield a strong VCD couplet that reflects molecular structures. Use of VCD couplets for biomacromolecular structural studies has been hampered by severe signal overlap caused by numerous functional groups that originally exist in biomacromolecules. Nitrile, isonitrile, alkyne, and azido groups show characteristic IR absorption in the 2300-2000 cm^{-1} region, where biomolecules do not strongly absorb. We herein examined the usefulness of these functional groups as chromophores to observe a strong VCD couplet that can be readily interpreted by theoretical calculations. Studies on a chiral binaphthyl scaffold possessing two identical chromophores showed that nitrile and isonitrile groups generate moderately strong but complex VCD signals due to anharmonic contributions. The nature of their anharmonic VCD patterns is discussed by comparison with the VCD spectrum of a mono-chromophoric molecule and by anharmonic DFT calculations. On the other hand, through studies on a diazido binaphthyl and a diazido monosaccharide, we demonstrated that azido group is more promising for structural analysis of larger molecules due to its simple, strong VCD couplet whose spectral patterns are readily predicted by harmonic DFT calculations.

2.2 Introduction

Through-space interactions of two suitable chromophores yield an exciton circular dichroism couplet, i.e., a pair of positive and negative Cotton effects. The sign and intensity of the couplet reflects the absolute arrangement of the electric dipole transition moments (EDTMs) associated with the chromophores. This phenomenon has provided a basis for a structural analysis method using electronic circular dichroism (ECD) spectroscopy, named the exciton chirality method.[1-4] With proper understanding on EDTMs and the conformations of analyte molecules, the ECD exciton chirality method has revealed the structures of various small molecules[3, 5, 6] as well as medium-sized molecules (500 to 2,000 Da).[3, 7-9] Use of this method has often been associated with introduction of chromophores when the original molecules do not show informative or strong ECD signals. A repertoire of UV chromophores with various absorption coefficients and absorption wavelength such as para-substituted benzoyl groups have been applied.[2-4] Chromophores with a longer absorption wavelength (e.g., para-porphyrin-substituted benzoyl groups at 420 nm) are also available to extract structural information from molecules that are originally ECD-active to a spectrally transparent region, where the original molecules do not show informative signals.[10,11] Such chromophores have also been used to enhance ECD signals for submicromolar analysis.[3, 4] Careful use of chromophores with longer absorption maxima has enabled the analysis of much bigger molecules (e.g., biomacromolecules, polymers, and multimolecular systems).[12-15] However, such bulky, hydrophobic, π conjugated chromophores are less favourable for studying the native structures of biomacromolecules. Their bulkiness and complex electronic natures also significantly increase the computational cost when theoretical calculations are used to interpret exciton couplet.

Meanwhile, IR chromophores, which are much smaller and thus less perturb biomolecular structures, have been known to generate a vibrational circular dichroism (VCD) couplet. This phenomenon was originally studied under the name of coupled oscillator.[16] We reinvestigated VCD couplets stemming from C=O stretching vibrations of a series of small molecules and reported a correlation between the sign of the couplet and the dihedral angle θ of two carbonyl groups.[17, 18] On the basis of this correlation, structures of various molecules have been analysed with or without introduction of extra carbonyl chromophores, in a similar manner to the applications of the ECD exciton chirality method.[19-30] VCD couplet is also used for signal augmentation that leads to VCD measurements with reduced sample amount and with a short acquisition time. For example, VCD couplet of **1** was observed even at a concentration of 2.5 mM.[17] Such a VCD couplet approach obviously is not applicable to molecules with almost parallel or antiparallel carbonyl groups and should be used with care for flexible molecules and multichromophoric systems as is the case for the ECD exciton chirality method. Theoretical studies by other groups suggested that such VCD couplets are only partially due to excitonic interactions and warned about possible exceptions to this correlation. [31-34] Although such exceptions have not yet been experimentally witnessed for C=O vibrations, use of computations such as density functional theory (DFT) to interpret VCD couplets is highly recommended.

Structures and intermolecular interactions of biomacromolecules in the solution state have been studied using multiple analytical methods including ECD spectroscopy, fluorescence resonance energy transfer (FRET), and electron spin labels. The VCD couplet approach should serve as a new analysis tool with less physicochemical perturbation. Because biomacromolecules possess many carbonyl groups, extraction of their structural information needs introduction of IR chromophores with less overlap with their intrinsic IR absorption. ^{13}C isotope labelling has been used for this purpose to red-

shift the C=O stretching frequency by ca. 40 cm^{-1}.[35] However, this shift is not sufficient to isolate the VCD signals without any overlap with $^{12}C=O$ and other vibrational signals. VCD couplets originating from other chromophores such as C=S, C-O, C=N and C-H have also been studied.[32, 33, 36-38] Yet to be examined are chromophores in the 2300-2000 cm^{-1} region, where the fundamental vibrational modes of common biomacromolecular functional groups do not exist.

Chromophores with X≡Y and X=Y=Z systems such as alkyne, nitrile, isonitrile, azido, allene, and carbodiimide exhibit characteristic IR absorptions in a spectrally transparent region for biomacromolecules. Some of these chromophores have been utilized as vibrational probes in biomacromolecular IR studies (e.g., oligonucleotide conformations, β-amyloid aggregation, and drug-protein interactions).[39-41] In the present work, we have studied the VCD spectra of **2-6** to test the applicability of -CN, -NC, -C≡CH, and $-N_3$ groups as a chromophore for observing their VCD couplet **(Figure 2.1)**. Dicarbonyl **1** was also studied to show a typical VCD couplet. Allene and carbodiimide were excluded because of their axially chiral structures that yield strong intrinsic VCD signals on their own.[42-44] The core structure 2,2′-disubstituted-1,1′-binaphthalene was chosen as a C2 symmetrical, rigid scaffold to regulate the distance (*r*) and the dihedral angle (θ) of the two chromophores. We aim for finding chromophores showing a strong VCD couplet whose shape is reliably predicted by theoretical calculations. An expected obstacle to this goal is that anharmonic vibrational contribution may be significant in the 2300-2000 cm^{-1} region and that its influence on the VCD spectra of chosen chromophores has not been studied. While this study focuses on small model molecules, the insight obtained here can be transferred to study larger molecules. Here, with using harmonic and anharmonic DFT calculations, we discuss the usefulness of these chromophores as well as their anharmonic behaviours.

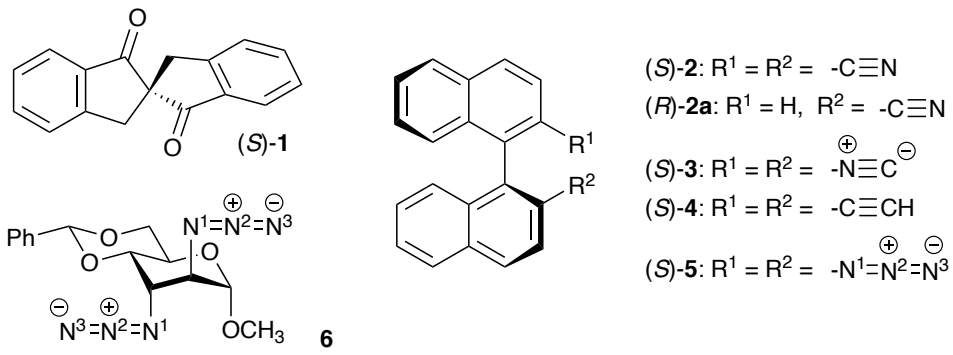


Figure 2.1: Structures of **1-6**. Note that the stereochemical nomenclature becomes *R* for **2**a.

2.3 Experimental and computational details

2.3.1 Experimental details

Compounds **2, 2a**, and **4** were prepared as a racemate starting from (±)-2,2′-dibromo-1,1′-binaphthyl and then enantioseparated using chiral column. Both enantiomers of 3 and 5 were synthesized from (*S*)- and (*R*)-1,1′-binaphthyl-2,2′-diamine. Compound **6** was synthesized from methyl α-D-glucopyranoside (see 2.5 Experimental section). For vibrational spectroscopy, each sample was dissolved in $CHCl_3$ or $CDCl_3$ and placed in a 50-μm or 100-μm BaF_2 cell. VCD and IR spectra were recorded using a JASCO FVS-6000 spectrometer with 4 cm^{-1} resolution for 1500 and 16 scans, respectively. Spectra in the 2300-2000 cm^{-1} region were measured using an InSb detector and an optical filter that passes through 2400-1900 cm^{-1} light, while those in the region below 1900 cm^{-1} were measured using an MCT detector with an optical filter that passes through light lower than 2200 cm^{-1}. The modulation frequency of the photoelastic modulator was set to 2127 cm^{-1} (for the measurement of 2300-2000 cm^{-1}) or 1350 cm^{-1} (for the measurement below 1900 cm^{-1}). With these spectrometer settings, each sample

showed a virtually mirror-image VCD spectrum to its enantiomer in the 2300-2000 cm^{-1} region (**Figure S1**). To obtain more accurate spectra, the VCD spectra of *S* enantiomers of **1-5** were corrected by their *R* enantiomers using the following equation.

$$\Delta\varepsilon(S)_{corrected} = 1/2[\Delta\varepsilon(S)_{raw} - \Delta\varepsilon(R)_{raw}]$$

The VCD spectrum of **6** and all the IR spectra were corrected by solvent spectra obtained under the identical measurement conditions.

2.3.2 Computational details

DFT calculations were carried out on Gaussian 16 package.[45] Initial conformational search was carried out using molecular mechanics on Spartan 18 program.[46] Obtained geometries within 20 kJ/mol from the most stable were submitted to DFT optimization at B3PW91/6-311++G(d,p) level of theory, which predicted only one conformer for each compound within 2.0 kcal/mol energy window. For these geometries, IR and VCD spectra were calculated using the same level of theory. The calculated frequencies, dipole strength, and rotational strength were converted to IR and VCD spectra on GaussView 6 software using a peak half-width at half height of 6 cm^{-1}. Anharmonic calculations of **2**, **2a**, and **3** were performed on a default program implemented in Gaussian 16. Calculated frequencies were scaled using factors mentioned in figure captions.

2.4 Results and discussion

2.4.1 Conformational analysis and harmonic vibrational calculations

To study the relative orientation of the two chromophores in **1-5**, MMFF conformational search and the following structural optimization at DFT/B3PW91/6-311G++(d,p) were carried out for their (*S*)-enantiomers. Only one stable conformer each existed for **2-4** due to the rigidity of the binaphthalene scaffold and the *sp* hybridized linear shape of nitrile, isonitrile, and alkyne (**Figure 2.2a**). For the calculations of **2**, chloroform molecules coordinating to nitrile groups are explicitly included as this treatment provided a better VCD agreement in the case of anharmonic calculations (vide infra). Dicarbonyl model compound **1** also displayed only one conformer. Interestingly, compound **5** also showed only one conformer within a 2.0 kcal/mol energy window despite the rotatable σ bonds between the naphthalene moiety and the chromophoric substituents. See supporting figures for computational results of less stable conformers of **5**.

Using these conformers, *r* and θ of **1-5** were analysed. This work defined these values based on the ideal localization of the vibrational motions of these chromophores. Namely, *r* was defined as the distance between two centres of the masses of atoms consisting of C=O bond (for **1**), each triple bond (for **2-4**), or of azido (for **5**), whereas θ was defined as the dihedral angle made by four atoms consisting of two C=O bonds (for **1**) or triple bonds (for **2-4**), or by the four terminal nitrogen atoms N3, N1, N1′, and N3′ (for **5**). Except for **1** and **5**, two chromophores in these structures were oriented with *r* and θ values of ca. 4 Å and +90°, respectively (**Figure 2.2a** and **Table 1.1**).

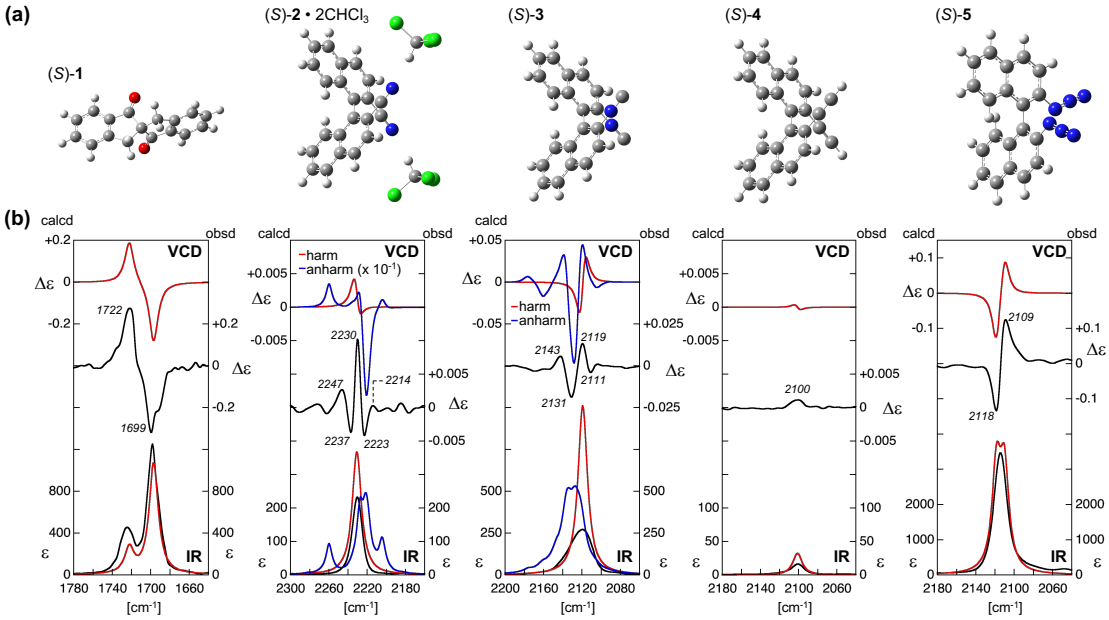


Figure 2.2: (a) Most stable conformers of (*S*)-**1**, (*S*)-**2**•2$CHCl_3$, (*S*)-**3**, (*S*)-**4**, and (*S*)-**5**. (b) VCD (top) and IR (bottom) of measured (black) and calculated (red for harmonic DFT and blue for anharmonic DFT) spectra. Wavenumbers at the extrema of observed VCD signals are labelled in italic. Measurement conditions: *c* 0.04 M in $CDCl_3$ (for 1), 0.2 M in $CHCl_3$ (for **2** and **3**), 0.6 M in $CHCl_3$ (for **4**), or 0.04 M in $CHCl_3$ (for **5**); *l* 50 μm. Calculation conditions: B3PW91/6-311++G(d,p). Frequency scaling factors: 0.953 (for **1**), 0.951 (for **2**, harmonic), 0.967 (for **2**, anharmonic), 0.966 (for **3**, harmonic), 0.987 (for **3**, anharmonic), 0.952 (for **4**), and 0.930 (for **5**).

Fundamental vibrational transitions of the most stable conformers of **1-5** were calculated using harmonic DFT calculations at the B3PW91/6-311++G(d,p) level of theory. Two fundamental vibrational modes were predicted in the 2300-2000 cm^{-1} region for each **2-5**. Corresponding C=O stretching modes of **1** were predicted at around 1700 cm^{-1} (scaled frequency). The dipole strength D_n and the rotational strength R_n associated with these vibrational modes are summarized in **Table 1** and the simulated VCD and IR spectra based on these values are shown in **Figure 2.2b**. Among **1-5**, D_n and R_n values of dialkyne **4** were 10 to 100 times smaller compared to those of the other molecules, as expected from the non-polar nature of -C≡CH group, which leads to very weak IR absorption for C≡C stretching. This indicated experimental difficulties in observing a VCD couplet of **4**. R_n of **1, 2, 3,** and **5** were predicted to be in the same order of magnitude despite the differences in the values of *r*, θ, and D_n. Nevertheless, due to the small

frequency differences of the two harmonic modes ($\nu_2 - \nu_1$) of **2** and **3**, their positive and negative components were predicted to mostly compensate each other. The frequency of the two azido vibrational modes of **5** differ by 8 cm^{-1}, which gave rise to a predicted strong VCD couplet.

Table 0.1: Interchromophoric distance (*r*) and dihedral angle of two chromophores (θ) of the most stable conformers of **1-6**, and predicted scaled frequencies (ν_n), dipole strengths (D_n), and rotational strengths (R_n) for the stretching vibrations of the discussed chromophores calculated at harmonic DFT/B3PW91/6-311++G(d,p).

	(*S*)-**1**	(*S*)-**2** • 2$CHCl_3$	(*S*)-**3**	(*S*)-**4**	(*S*)-**5**	(*R*)-**2a** • $CHCl_3$	**6**
r [Å]	3.2	4.3	4.2	4.4	5.3	–	5.6
θ [°]	-116	+97	+95	+94	-23	–	+122
ν_1 [cm^{-1}]	1696.9	2230.7	2118.7	2101.1	2110.2	2227.6	2107.1
D_1 [10^{-40} esu^2 cm^2]	1034	120	424	11.7	2124	160	1846
R_1 [10^{-44} esu^2 cm^2]	-706	-239	+537	+23.6	+260	+2.3	+378
ν_2 [cm^{-1}]	1722.2	2230.9	2119.8	2101.3	2118.4	–	2118.9
D_2 [10^{-40} esu^2 cm^2]	232	152	381	13.7	2169	–	558
R_2 [10^{-44} esu^2 cm^2]	+483	+243	-545	-25.0	-316	–	-558

2.4.2 Experimental spectra of dinitrile, diisonitrile, and dialkyne compounds

VCD spectra in the 2300-2000 cm^{-1} region have rarely been documented, especially since the development of Fourier-transform (FT) VCD spectrometers.[44, 47-50] To achieve a high S/N ratio in this region, we inserted an optical filter that passes through 2400-1900 cm^{-1} light to a FT-VCD spectrometer and used an InSb detector. Furthermore, to maximize the reliability of the observed spectra, both enantiomers of **2-5** were prepared and the VCD spectra of the enantiomeric pairs were compared. Each enantiomeric pair showed almost mirror-image VCD patterns in the 2300-2000 cm^{-1} region (**Figure S1**), which validated the accuracy of the VCD measurement in this region. The following

spectral analysis is based on the "enantiomer-corrected" VCD spectra of (*S*)-enantiomers of **1-5** (**Figure 2.2b**).

Two nitrile groups (**2**) and isonitrile groups (**3**) gave rise to VCD signals with Δε of an order of 0.01, while two azido groups (**5**) and carbonyl groups (**1**) generated 10- to 100-fold stronger VCD absorption. Meanwhile, **4** showed a negligibly small VCD band with very small IR absorption (Δε 0.001 and ε 16 at 2100 cm^{-1}), consistent with the calculated small D_n and R_n values. Thus, we concluded that the terminal C≡CH group is not suitable for sensitively observing VCD couplet.

A VCD couplet originating from C=O stretching vibrations is presented by (*S*)-**1** (Δε -0.31 at 1699 cm-1 and Δε +0.27 at 1722 cm-1). In this case the IR absorption bands also split. Its negative-positive VCD couplet was well reproduced by harmonic DFT calculations (**Figure 2.2b**). In contrast, the observed VCD of dinitrile **2** and diisonitrile **3** were more complex and showed a greater number of peaks compared to the harmonically calculated simple VCD spectra. With alternating positive and negative patterns of the observed VCD spectra, reliable interpretation of these VCD spectra in terms of VCD couplet is not possible. Similarly, attempts to correspond the harmonically predicted VCD signals of **2** and **3** to two of the observed VCD peaks are not appropriate. Previous VCD studies on deuterated molecules reported somewhat similar VCD patterns in the 2300-2000 cm^{-1} region and ascribed these to anharmonic VCD signals. [48, 49] Indeed, the IR band of **3** at 2119 cm^{-1} showed a wider skirt on the higher frequency side, which indicated involvement of anharmonic signals. Dinitrile **2** yielded a more symmetrical IR band, but nitrile groups are also known to exhibit anharmonic signatures in biomolecular IR studies.[51] Therefore, the complex, yet strong VCD features of **2** and **3** were ascribed to anharmonic VCD signals involving -CN and -NC vibrations, respectively.

2.4.3 Analysis of anharmonic VCD features of dinitrile and diisonitrile compounds

As mentioned earlier, this work aims for identifying chromophores exhibiting a strong VCD couplet that can be interpreted by theoretical calculations toward its applications to structural analysis of biomacromolecules. Nitrile and isonitrile groups showed strong VCD signals, but the complex signal patterns seemed disadvantageous for structural analysis. Nonetheless, the usefulness of nitrile functionality as an IR reporter in biochemistry and the scarcity of insight into anharmonic VCD in the 2300-2000 cm^{-1} region prompted us to further investigate the VCD of **2** and **3** regarding on the following two questions:

(1) Harmonic VCD signals are known to be augmented by interactions of two IR chromophores, but are anharmonic VCD signals also augmented by such interactions?

(2) Are these anharmonic VCD spectra still useful for extracting structural information of the analyte molecule by using anharmonic DFT calculations?

To address the first question, mononitrile (*R*)-**2a** was synthesized to analyse its VCD spectrum in comparison with that of dinitrile (*S)*-**2**. Mononitrile **2a** was prepared as a racemate and then enantioseparated using CHIRALPAK® IB column, which led to the first-eluted (*S*)-(−)-**2a** and the second-eluted (*R*)-(+)-**2a** (**Figure 2.3a, Figure S3**).[52] Enantiopure (*R*)-**2a** showed a twice weaker IR band at 2228 cm^{-1} than that of dinitrile (*S*)-**2** (**Figure 2.3**). Meanwhile, no discernible VCD signals were observed for (*R*)-**2a**. Both harmonic and anharmonic (vide infra) DFT calculations of (*R*)-**2a**•$CHCl_3$ (**Figure 2.3c**) predicted ca. 100-fold smaller VCD signals compared to those calculated for **2** (**Table 2.1** and **Figure 2.3b**). These results, for the first time, experimentally verified that interactions of chromophores enhance anharmonic VCD intensity.

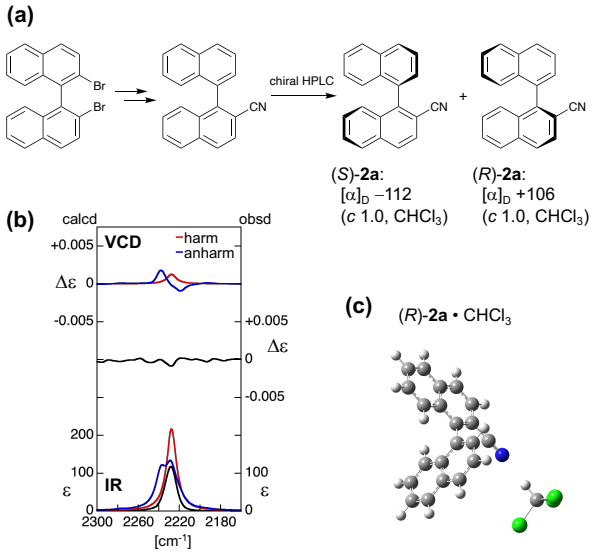


Figure 2.1: (a) Preparation of enantiomerically pure **2a**. (b) VCD (top) and IR (bottom) of measured (black) and calculated (red for harmonic DFT and blue for anharmonic DFT) spectra of (*R*)-**2a**. (c) Most stable conformer of (*R*)-**2a**•$CHCl_3$. Measurement conditions: *c* 0.2 M in $CHCl_3$, *l* 100 μm, VCD spectrum of (*R*)-(+)-**2a** was corrected by that of (*S*)-(-)-**2a**. Calculation conditions: B3PW91/6-311++G(d,p). Frequency scaling factors: 0.950 (for harmonic) and 0.967 (for anharmonic).

Regarding on the second question, because the complex VCD signals of **2** and **3** were generated by interplays of two chromophores, the spectral patterns should reflect the orientation of chromophores as well as their local environment. Interpretation of anharmonic VCD spectra by means of anharmonic calculations has been reported, [53, 54] but their utility in the 2300-2000 cm^{-1} region has not been studied. Aiming at correlating anharmonic VCD signals with molecular structures, we performed VCD calculations of (*S*)-**2** and (*S*)-**3** using anharmonic DFT settings implemented in Gaussian 16 program. Among several functionals and basis sets tested, B3PW91/6-311++G(d,p) level well reproduced the observed VCD features of **3** (**Figure 2.2**). Namely, experimentally observed small negative band at 2111 cm^{-1}, strong positive band at 2119 cm^{-1}, strong negative band at 2131 cm^{-1}, and moderately strong positive band at 2143 cm^{-1} were all seen in the calculated spectrum with similar shape and intensity. Application of the same level of theory to (*S*)-**2** with two explicit chloroform molecules resulted in a VCD

spectrum partially like the observed one. The observed positive VCD peak at 2230 cm^{-1} and negative one at 2237 cm^{-1} may have cancelled out in the anharmonically predicted VCD spectrum. Comparison of the observed spectra and anharmonically calculated ones in the region below 1900 cm^{-1} also showed a moderate agreement (**Figure S4**). Overall, these preliminary results showed promise for deducing molecular structures from anharmonic VCD signals.

These results also highlighted obstacles to be overcome for studying biomacromolecules using these chromophores. First, the high computational cost for anharmonic DFT (ca. 100 times longer computational time than harmonic DFT) hampers the VCD calculations of many candidate structures. Second, solvent effects may be properly included into VCD calculations to well reproduce observed spectral patterns.[55] In the case of dinitrile **2**, solvent effects (e.g., implicit, and explicit) drastically changed predicted anharmonic VCD results (**Figure S5**). Last, the dependence of the predicted VCD shape to the basis set and functional of a given system is yet to be investigated. See **Figure S5** for anharmonic VCD spectra of **2** predicted by other calculation conditions. Although investigation of the second and last points are in progress by our group, chromophores with less anharmonic nature seem more practical.

2.4.4 VCD couplet by diazido compounds

Unlike nitrile and isonitrile chromophores, the interaction between two azido groups in (*S*)-**5** produced a rather simple positive-negative VCD couplet. This observation was in line with the previous insight that accidental Fermi resonance is less obvious for azido group.[40, 56] Moreover, the observed VCD intensities (Δε +0.12 at 2109 cm^{-1} and Δε -0.13 at 2118 cm^{-1}) are one order higher than those of nitrile and isonitrile (**Figure 2.2**). This positive-negative couplet was accurately reproduced by harmonic VCD calculations of (S)-**5**. Thus, azido group fulfils the requirements as VCD chromophores to extract structural information of biomacromolecules.

To test the applicability of azido group to biomolecules, we prepared a diazido monosaccharide derivative 6, whose azido groups are connected to sp^3 aliphatic carbons. The experimental VCD spectrum of **6** in $CHCl_3$ showed a strong positive-negative couplet without any obvious anharmonic VCD signals (Δε +0.10 at 2104 cm^{-1} and Δε -0.034 at 2119 cm^{-1}) (**Figure 2.4**). Such a positive-negative couplet of 6 was also observed in a DMSO-h_6 – H_2O (9:1) mixed solvent (**Figure S6**). We also carried out MMFF conformational search and the following DFT structural optimization to investigate the orientation of the azido groups in **6**. Despite the rotatable nature of the σ bonds between the azido group and the pyranose ring, only one conformer with r of 5.6 Å and θ of +122° was found within a 3.0 kcal/mol energy window at B3PW91/6-311++G(d,p) (**Figure 2.4b**). The second most stable conformer with ΔE = 3.12 kcal/mol differed only in the rotation of C1 methoxy group (**Figure S2b**). Harmonic VCD calculations of the most stable conformer well reproduced the positive-negative couplet (**Figure 2.4a**), which supported the applicability of azido chromophores to structural analysis of biomolecules. Spectral agreement between the observed and the calculated was also seen in the region below 1600 cm^{-1} (**Figure S7a**). Azido VCD couplets are reliably predicted also with using

a smaller level of theory (B3PW91/6-31G(d)) as shown in **Figure S7b**. Another practical advantage of azido is that protocols for its incorporation to various biomolecules have been established and some of azido-containing biomolecules (e.g. amino acids and sugars) are commercially available owing to the development of bioorthogonal alkyne-azide chemistry.[57] Application of azido groups to more complex molecular systems should be reported in due course.

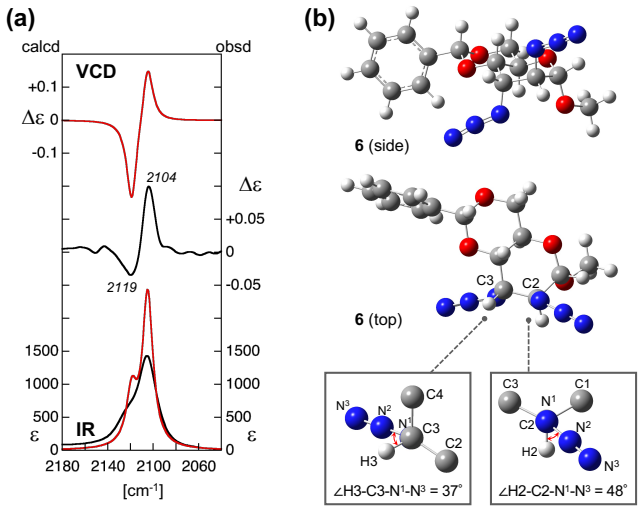


Figure 3.2: (a) VCD (top) and IR (bottom) of measured (black) and calculated (red) spectra of **6**. (b) The most stable conformer of **6**. Newman projections for C2-N and C3-N bonds are shown at the bottom, in which dihedral angles made by azido group and methine C-H are illustrated by red arrows. Measurement conditions: *c* 0.04 M in $CHCl_3$, *l* 50 μm. Calculation conditions: harmonic DFT at B3PW91/6-311++G(d,p). Frequency scaling factor: 0.927.

Clockwise orientation of two azido groups of **6** showed a positive-negative VCD couplet, while anticlockwise **5** with a small dihedral angle (θ –23°) also exhibited a positive-negative couplet. Thus, azido VCD couplets of **5** does not follow a coupled oscillator model.[16] The stretching vibrations of two azido groups in both **5** and **6** were predicted to be out-of-phase for the lower frequency transition and in-phase for the higher frequency one. To obtain further insight into the relationship between the chromophoric orientation and the signal pattern, we calculated the VCD spectra of two azidomethane

(CH_3-N_3) molecules and two azidobenzene (Ph-N_3) molecules. VCD computations of these model systems with 10° incremental changes of θ showed a clear dependence of VCD pattern to θ (**Figure 2.5**). For example, the intensity of the couplet becomes smaller when θ is close to 0° and 180°. These results indicated that the EDTM of azido vibrational modes may be approximated to lie on the line connecting N1 and N3. Azidomethane molecules and azidobenzene molecules each placed with θ of ca. –23° showed a negative-positive VCD couplet. The discrepancy between the VCD properties of these bimolecular computational systems and those of **5** suggested that the VCD signals generated from N=N=N stretching vibration is readily perturbed by molecular structures other than azido moieties. Thus, we recommend to always use theoretical calculations for interpreting azido VCD couplets.

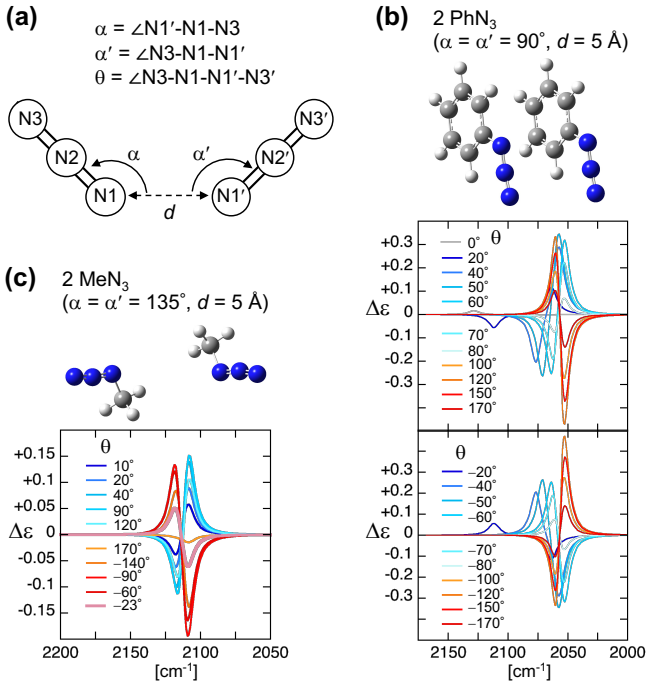


Figure 3.3: Calculated VCD spectra of two simple azides as a function of the dihedral angle θ with fixed values of α, α′, and d. (a) Definitions of the values α, α′, d, and θ in this work. (b) Calculated VCD spectra of two azidobenzene molecules. Two phenyl groups are in cis orientation. (c) Calculated VCD spectra of two azidomethane molecules. Two methyl groups are in trans orientation. For both azidobenzene and azidomethane, structures with θ = 0° are drawn. Calculation conditions: B3PW91/6-311++G(d,p).

2.5 Conclusion

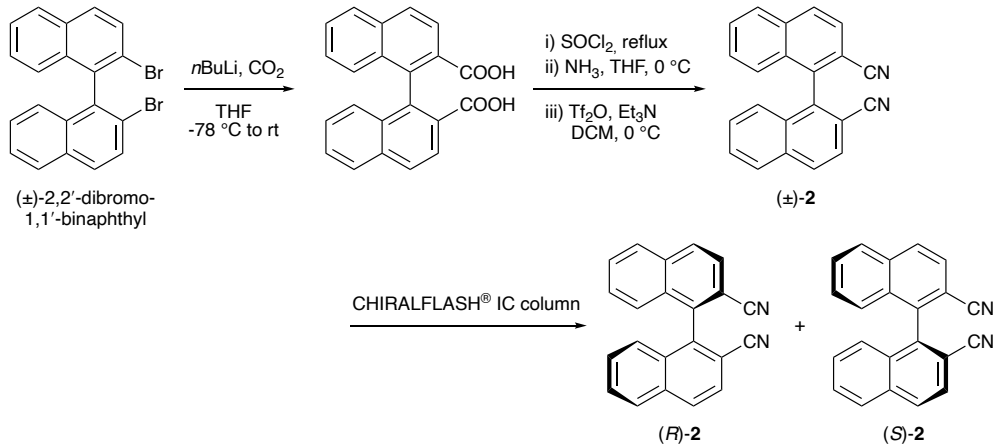
VCD spectroscopy is a useful technique to analyse the structure of biomolecules, but its use for biomacromolecules has been limited due to the low signal intensity and severe signal overlap. This study demonstrated that introduction of chromophores in the 2300-2000 cm^{-1} region yields strong VCD signals that reflect molecular structures. Nitrile and isonitrile groups exhibited strong but complex anharmonic VCD signals. This work revealed that interplays between two chromophores enhance anharmonic VCD intensities. We also demonstrated a possibility that complex anharmonic VCD signals can be interpreted by means of anharmonic DFT calculations. Azido group is more advantageous due to its stronger VCD couplet whose pattern can be readily predicted by harmonic DFT calculations. In combination with other techniques (e.g., 2D IR, FRET and ESR), VCD spectroscopy using these chromophores should facilitate future structural studies of biomacromolecules in the solution

2.6 Experimental Section: Synthesis of 2, 3, 4, 5, 2a, and 6

2.6.1 General Procedures

^{1}H NMR (500 MHz) and ^{13}C NMR (125 MHz) spectra were recorded on a Varian Inova instrument at 25 °C. Chemical shift values (δ) are reported in ppm relative to tetramethylsilane. The following abbreviations were used for signal multiplicities: s = singlet; d = doublet; m = multiplet. Optical rotations were measured on a JASCO P-1020 polarimeter at the sodium D-line under ambient temperature, and reported as $[\alpha]_D$ (concentration in grams/100 mL solvent). Enantioseparation of **2** and **4** was performed by using a uf-3020SZB2 pump (Denso Sangyo, Japan) equipped with a Shimamura (Japan) YRU-880 midget UV-RI detector, using a Daicel CHIRALFLASH® IC column (3.0 cm ϕ × 10 cm). Enantioseparation of **2a** was conducted on two JASCO PU-2086 intelligent pumps equipped with a JASCO MX-2080-31 solvent mixing module on a PU-2075 intelligent UV/Vis detector, using a Daicel CHIRALPAK® IB column (1.0 cm ϕ × 25 cm). Spectroscopic grade $CHCl_3$ and $CDCl_3$ were purchased from Wako Pure Chemical Industries and Cambridge Isotope Laboratories, respectively. Purchased chemicals were used without further purification. Synthesis of (*S*)-**1** and (*R*)-**1** is reported previously.[17]

2.6.2 Synthesis of (*S*)-2 and (*R*)-2

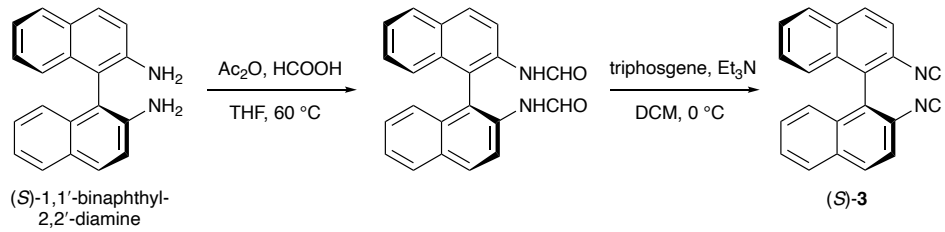


from (±)-2,2′-dibromo-1,1′-binaphthyl to (±)-**2** was carried out in a similar manner to a reported procedure.[58] (±)-2,2′-dibromo-1,1′-binaphthyl (206 mg, 499 μmol) in THF (10 mL) was added *n*-butyllithium in hexane (2.5 M, 0.48 mL) dropwise under N_2 at -78 °C. After 30 mins, N_2 was replaced with CO_2, and the mixture was stirred at -78 °C for 1.5 h and then gradually brought to rt overnight. The reaction was quenched with 2 M HCl aq and then extracted using Et_2O. The organic layer was washed with brine, and then dried over $MgSO_4$. After removal of the solvent, the mixture was purified by silica-gel column chromatography (hexane-EtOAc-AcOH = 95:5:1 to 80:20:1), which afforded (±)-1,1′-binaphthyl-2,2′-dicarboxylic acid[59] (105 mg, 61%).

The diacid (105 mg) was added $SOCl_2$ (2 mL) and the mixture was refluxed. After 15 mins, the mixture was dried under reduced pressure. The residue was dissolved in THF (1 mL) and then bubbled with NH_3 gas at 0 °C for 10 mins. After removal of the solvent, the residue was dissolved in DCM (1 mL) and added Tf_2O (0.25 mL) and Et_3N (0.5 mL) and the mixture was stirred at 0 °C. After 40 mins, the mixture was diluted with EtOAc, washed sequentially with 2 M HCl aq and brine, and then dried over $MgSO_4$. After removal of the solvent, the mixture was purified by silica-gel column chromatography (hexane-EtOAc = 8:1 to 1:1), which afforded (±)-**2**[60] (34 mg, 37% from diacid).

Enantioseparation of (±)-**2** was carried out on a CHIRALFLASH® IC column (3.0 cm ϕ × 10 cm) using hexane-EtOAc = 4:1, which led to the first-eluted (*R*)-(-)-**2** and the second-eluted (*S*)-(+)-**2**. (*R*)-(-)-**2**: $[\alpha]_D$ -43.8 (*c* 1.0, $CHCl_3$); lit[61] $[\alpha]_D$ -70.5 (*c* 1, $CHCl_3$). (*S*)-(+)-**2**: $[\alpha]_D$ +44.3 (*c* 1.0, $CHCl_3$); lit[61] $[\alpha]_D$ +68.6 (*c* 1.54, $CHCl_3$).

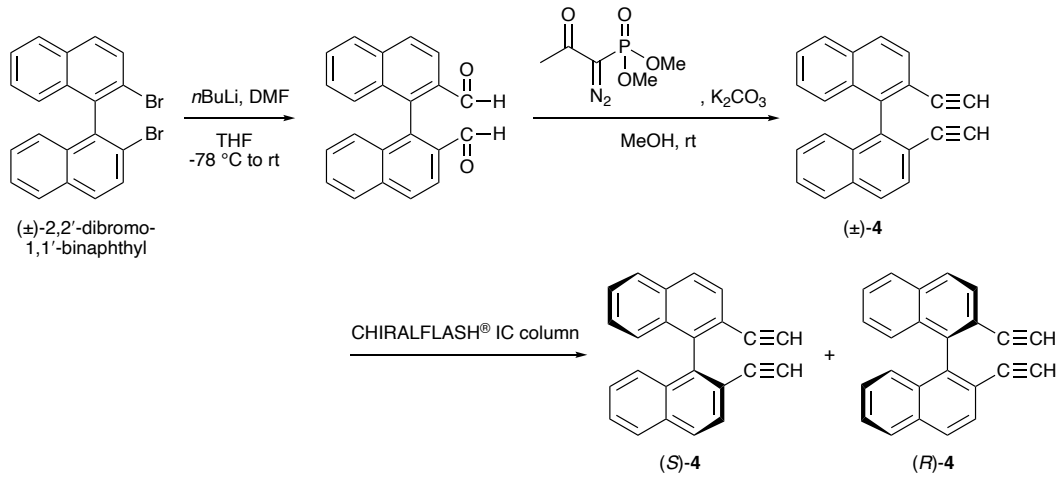
2.6.3 Synthesis of (*S*)-3 and (*R*)-3



(*S*)-**3** was synthesized in a similar manner to a reported procedure.[62] Formic acid (500 μL) was added to Ac_2O (60 μL) at 0 °C and the solution was stirred at 60 °C for 30 mins. To this solution, THF (0.5 mL) and (*S*)-1,1′-binaphthyl-2,2′-diamine (32 mg, 113 μmol) in THF (0.5 mL) were successively added at 0 °C and the mixture was stirred at 60 °C. After 3 h, the mixture was diluted with EtOAc, washed sequentially with sat $NaHCO_3$ aq and brine, and then dried over $MgSO_4$. After removal of the solvent, the mixture was purified by silica-gel column chromatography (hexane-EtOAc = 1:3 to 1:19), which afforded (*S*)-1,1′-binaphthyl-2,2′-diformamide containing some impurities (23 mg). Without further purification, the diformamide (23 mg) and Et_3N (150 μL) in DCM (0.7 mL) was added a 0.3 mL DCM solution of triphosgene (33 mg) dropwise at 0 °C and stirred at rt. After 1.5 h, the solvent was removed and the residue was purified by silica-gel column chromatography (hexane-EtOAc = 3:1 to 1:3), which afforded (*S*)-**3**[62] (10 mg, 29% from diamine).

(*R*)-**3**[63] was prepared from (*R*)-1,1′-binaphthyl-2,2′-diamine using a similar procedure.

2.6.4 Synthesis of (*S*)-4 and (*R*)-4



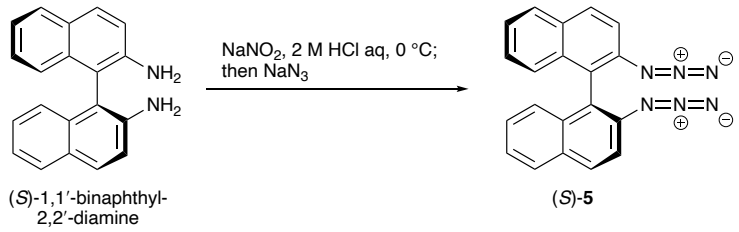
Conversion from (±)-2,2′-dibromo-1,1′-binaphthyl to (±)-**4** was carried out in a similar manner to a reported procedure.[7] (±)-2,2′-dibromo-1,1′-binaphthyl (418 mg, 1.01 mmol) in THF (10 mL) was added *n*-butyllithium in hexane (2.5 M, 0.96 mL) dropwise under N_2 at -78 °C. After 30 mins, the mixture was added DMF (2.3 mL), stirred at -78 °C for 40 mins, and then gradually brought to rt overnight. The reaction mixture was diluted with Et_2O and the organic layer was washed with sat NH_4Cl aq. The aqueous layer was extracted with Et_2O. All organic layers were combined and dried over $MgSO_4$. After removal of the solvent, the mixture was purified by silica-gel column chromatography (hexane-EtOAc = 10:1 to 4:1), which afforded (±)-2,2′-diformyl-1,1′-binaphthyl[64] (85 mg, 27%).

The dialdehyde (85 mg, 274 μmol) and K_2CO_3 (85 mg) in MeOH (10 mL) was added Bestmann-Ohira reagent (165 μL) and stirred at rt for overnight. The reaction mixture was diluted with Et_2O and the organic layer was washed with sat NH_4Cl aq. The aqueous layer was extracted with Et_2O. All organic layers were combined and dried over $MgSO_4$.

After removal of the solvent, the mixture was purified by silica-gel column chromatography (hexane-EtOAc = 25:1 to 20:1), which afforded (±)-**4**[64] (27 mg, 33%).

Enantioseparation of (±)-**4** was carried out on a CHIRALFLASH® IC column (3.0 cm ϕ × 10 cm) using hexane-EtOAc = 50:1, which led to the first-eluted (-)-**4** and the second-eluted (+)-**4**. The absolute configurations of these enantiomers were determined to be (*S*)-(-)-**4** and (*R*)-(+)-**4** by using VCD spectroscopy (Fig. S3a). (*S*)-(-)-**4**: $[\alpha]_D$ -28.4 (*c* 1.0, $CHCl_3$). (*R*)-(+)-**4**: $[\alpha]_D$ +27.2 (*c* 1.0, $CHCl_3$).

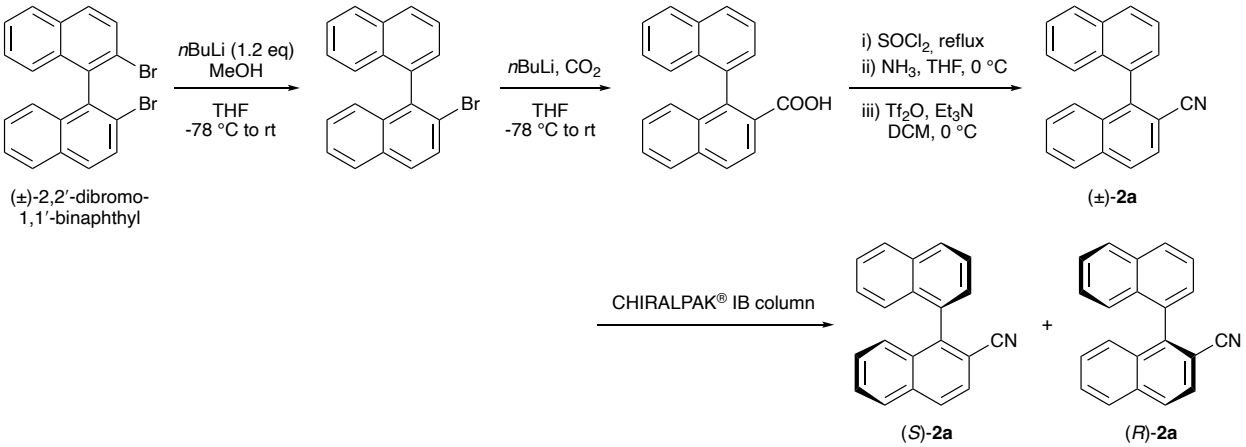
2.6.5 Synthesis of (*S*)-5 and (*R*)-5



(*S*)-**5** was synthesized in a similar manner to a reported procedure.[65] (*S*)-1,1′-binaphthyl-2,2′-diamine (142 mg, 499 μmol) in 2 M HCl aq (3.5 mL) was added $NaNO_2$ (105 mg) in H_2O (1 mL) and stirred at 0 °C. After 1 h, the mixture was added NaN_3 (135 mg) in H_2O (1 mL) dropwise at 0 °C and stirred at rt for overnight. The mixture was diluted with EtOAc, washed sequentially with sat $NaHCO_3$ aq and brine, and then dried over $MgSO_4$. After removal of the solvent, the mixture was purified by silica-gel column chromatography (hexane-EtOAc = 99:1), which afforded (*S*)-**5** (146 mg, 87%). Its NMR data were virtually identical with those reported for (±)-**5**.[65] (*S*)-(-)-**5**: $[\alpha]_D$ -42.1 (*c* 1.0, $CHCl_3$).

(*R*)-**5**[66] was prepared from (*R*)-1,1′-binaphthyl-2,2′-diamine using a similar procedure.

2.6.6 Synthesis of (*S*)-2a and (*R*)-2a



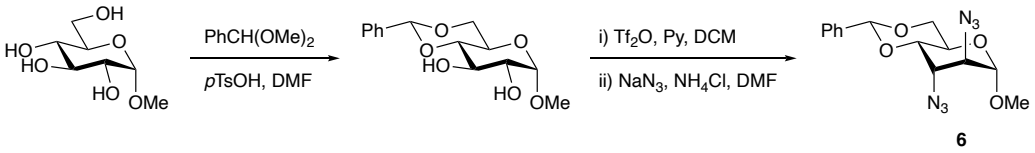
Conversion from (±)-2,2′-dibromo-1,1′-binaphthyl to (±)-1,1′-binaphthyl-2-carboxylic acid was carried out in a similar manner to a reported procedure.[66] (±)-2,2′-dibromo-1,1′-binaphthyl (412 mg, 1.00 mmol) in THF (10 mL) was added *n*-butyllithium in hexane (2.6 M, 0.46 mL) dropwise under N_2 at -78 °C. After 1 h, the mixture was added MeOH (1 mL) and stirred for 10 mins at -78 °C. The mixture was diluted with $CHCl_3$, washed with brine, and then dried over $MgSO_4$. After removal of the solvent, the residue was dissolved in THF (10 mL) and added *n*-butyllithium in hexane (2.6 M, 0.58 mL) dropwise under N_2 at -78 °C. After 15 mins, N_2 was replaced with CO_2, and the mixture was gradually brought to rt over 1 h. The reaction mixture was diluted with $CHCl_3$, washed with 2 M HCl aq, and then dried over $MgSO_4$. After removal of the solvent, the mixture was purified by silica-gel column chromatography ($CHCl_3$-MeOH = 15: 1), which afforded (±)-1,1′-binaphthyl-2-carboxylic acid[66] (165 mg, 55% in 2 steps).

The following conversion from monoacid to (±)-**2a** was performed in a similar manner to the synthesis of (±)-**2**. The monoacid (160 mg) was added $SOCl_2$ (3 mL), and the mixture was refluxed. After 10 mins, the mixture was dried under reduced pressure. The residue was dissolved in THF (2 mL) and then bubbled with NH_3 gas at 0 °C for 10 mins. After removal of the solvent, the residue was dissolved in DCM (2 mL) and added Tf_2O

(0.8 mL) and Et_3N (1 mL) and the mixture was stirred at 0 °C. After 40 mins, the mixture was diluted with EtOAc, washed sequentially with 2 M HCl aq, sat $NaHCO_3$ aq and brine, and then dried over $MgSO_4$. After removal of the solvent, the mixture was purified by silica-gel column chromatography (hexane-EtOAc = 15:1), which afforded (±)-**2a**[67] (80 mg, 54% from monoacid).

Enantioseparation of (±)-**2a** was carried out on a CHIRALPAK® IB column (1.0 cm □ × 25 cm) using hexane-EtOH = 95:5, which led to the first-eluted (-)-**2** at t_1 = 10.7 min and the second-eluted (*S*)-(+)-**2** at t_2 = 11.4 min. The separation factor α was calculated as 1.1, where t_0 = 4.0 min was used. The absolute configurations of these enantiomers were determined to be (*S*)-(-)-**2a** and (*R*)-(+)-**2a** by using VCD spectroscopy (Fig. S3b). (*S*)-(-)-**2a**: $[\alpha]_D$ -112 (*c* 1.0, $CHCl_3$). (*R*)-(+)-**2a**: $[\alpha]_D$ +106 (*c* 1.0, $CHCl_3$);[68] for (+)-**2a**: $[\alpha]_D$ +89.0 (*c* 1.0, $CHCl_3$).

2.6.7 Synthesis of 6



6

Methyl a-D-glucopyranoside (1.95 g, 10.0 mmol) in DMF (35 mL) was added benzaldehyde dimethoxy acetal (2 mL) and a catalytic amount of *p*-TsOH•H_2O and stirred at rt overnight. The mixture was diluted with $CHCl_3$, washed sequentially with sat $NaHCO_3$ aq and brine, and dried over $MgSO_4$. After removal of the solvent, the mixture was purified by recrystallization with EtOH and hexane, which afforded methyl 4,6-*O*-benzylidine-a-D-glucopyranoside (2.14g, 76%). Its ^1H NMR spectrum was virtually identical with a reported one.[69]

The 4,6-protected sugar (148 mg, 524 μmol) in DCM (3 mL) was added pyridine (900 μL) and trifluoromethanesulfonic anhydride (220 μL) and stirred at rt. After 30 mins, the mixture was diluted with DCM, washed sequentially with water and brine, and then dried over $MgSO_4$. After removal of the solvent, the residue was dissolved in DMF (10 mL), added NaN_3 (374 mg) and NH_4Cl (18 mg) and stirred overnight at 80°C. The mixture was brought to rt, diluted with EtOAc, washed sequentially with water and brine, and then dried over $MgSO_4$. After removal of the solvent, the mixture was purified by silica-gel column chromatography (hexane-EtOAc = 10:1 to 5:1), which afforded methyl 2,3-diazido-2,3-deoxy-4,6-*O*-benzylidine-α-D-altropyranoside **6**[70] (78 mg, 45% in 2 steps). ^{1}H NMR ($CDCl_3$) □ 7.51–7.47 (m, ArH, 2H), 7.41–7.34 (m, ArH, 3H), 5.62 (s, C<u>H</u>Ph, 1H), 4.66 (d, H-1, *J* = 0.8 Hz, 1H), 4.34–4.25 (m, H-6a, H-5, 2H), 4.10 (dd, H-3, *J* = 3.0, 3.0 Hz, 1H), 4.05 (dd, H-4, *J* = 9.3, 3.5 Hz, 1H), 3.83 (dd, H-2, *J* = 2.6, 1.0 Hz, 1H), 3.79 (dd, H-6b, *J* = 10.1, 10.1 Hz, 1H), 3.45 (s, OMe, 3H); ^{13}C NMR ($CDCl_3$) □ 137.0 (Ar), 129.5 (Ar), 128.6 (Ar), 126.3 (Ar), 102.5 (<u>C</u>HPh), 99.04 (C-1), 75.7 (C-4), 69.1 (C-6), 61.3 (C-2), 58.9 (C-5), 58.5 (C-3), 56.1 (OMe). $[\alpha]_D$ +20.8 (*c* 1.0, $CHCl_3$).

2.7 Supporting Information

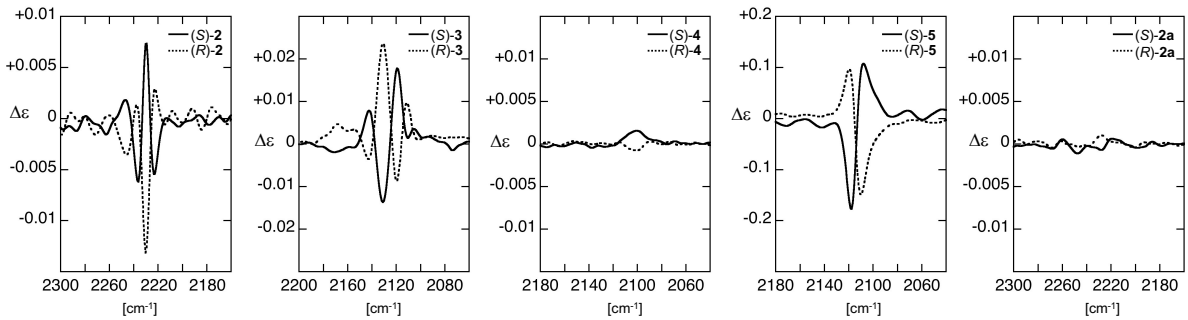
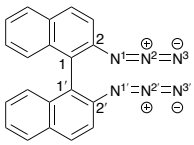


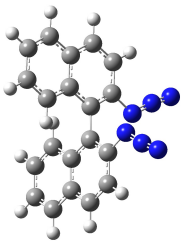
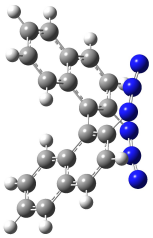
Figure S1: VCD spectra of enantiomeric pairs of **2-5** without "enantiomer correction" in the 2300-2000 cm^{-1} region. Measurement conditions are described in the caption of Figures **2** and **3**.

(a)

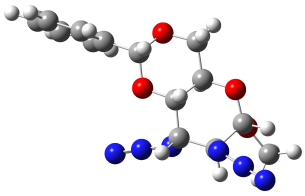


5

ϕ_1 = ∠C1-C2-N1-N2
ϕ_2 = ∠C1′-C2′-N1′-N2′
θ = ∠N3-N1-N1′-N3′

	 ϕ_1 = ca. 180°, ϕ_2 = ca. 180° (**conformer a**)	 ϕ_1 = ca. 0°, ϕ_2 = ca. 0° (**conformer b**)
B3PW91/6-311++G(d,p)	Δ*E* = 0.00 kcal/mol	converged to **conformer a**
B3PW91/6-31G(d)	Δ*E* = 0.00 kcal/mol	Δ*E* = 3.61 kcal/mol
B3LYP/6-311++G(d,p)	Δ*E* = 0.00 kcal/mol	converged to **conformer a**
B3LYP/6-31G(d)	Δ*E* = 0.00 kcal/mol	Δ*E* = 3.44 kcal/mol
M06-2X/6-31G(d)	Δ*E* = 0.00 kcal/mol	converged to **conformer a**

(b)



6 (less stable conformer)

Δ*E* = 3.12 kcal/mol
(B3PW91/6-311++G(d,p))

Figure S2: Less stable conformers of **5** and **6**. (a) Three geometries with (ϕ_1, ϕ_2) = (0°, 0°), (0°, 180°), and (180°, 180°) were submitted to DFT optimization. With B3PW91/6-311++G(d,p) level of optimization, all geometries were converged to a conformer with angles ϕ_1 and ϕ_2 of approximately 180° and 180° (θ = -23°). Some of different DFT levels yielded the second most stable conformer with ϕ_1 and ϕ_2 of approximately 0° and 0° with Δ*E* of higher than 3.0 kcal/mol. Conformers with (ϕ_1, ϕ_2) = (0°, 180°) could not be obtained by all the tested calculation conditions. (b) The second most stable conformer of 6 predicted at DFT/B3PW91/6-311++G(d,p).

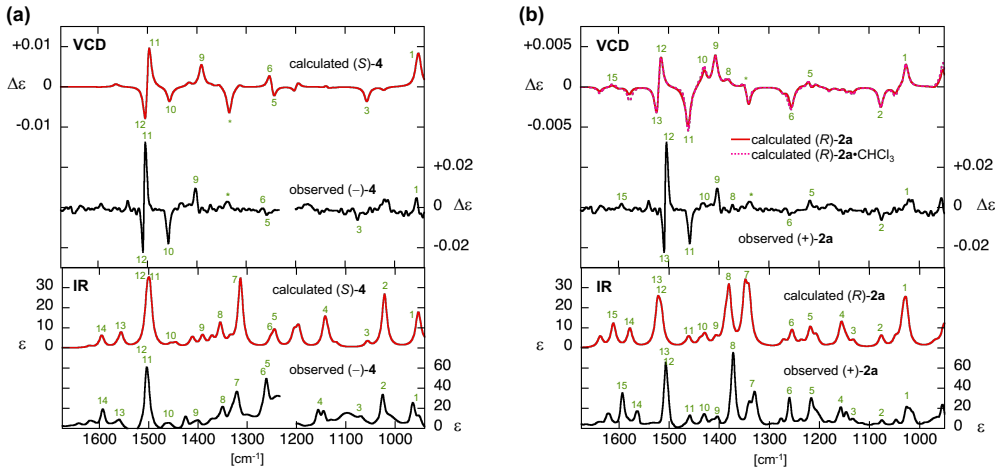


Figure S3: Assignment of the chirality of (a) (–)-**4** and (b) (+)-**2**a by comparison of calculated and observed VCD (top) and IR (bottom) spectra. Theoretical VCD spectrum for (*R*)-**2**a was almost superimposable to that for (*R*)-**2**a•$CHCl_3$. Part of corresponding VCD and IR peaks are labelled, while nonmatching peaks are indicated by asterisks. Measurement conditions: *c* 0.6 M in $CHCl_3$ (for (–)-**4**) or 0.3 M in $CDCl_3$ (for (+)-**2**a), *l* 50 μm (for (–)-**4**) or 100 μm (for (+)-**2**a), corrected by solvent spectra. Calculation conditions: B3PW91/6-311++G(d,p). Frequency scaling factor: 0.97 (for (*S*)-**4**) or 0.98 (for (*R*)-**2**a). Part of observed spectra are omitted due to strong solvent absorption.

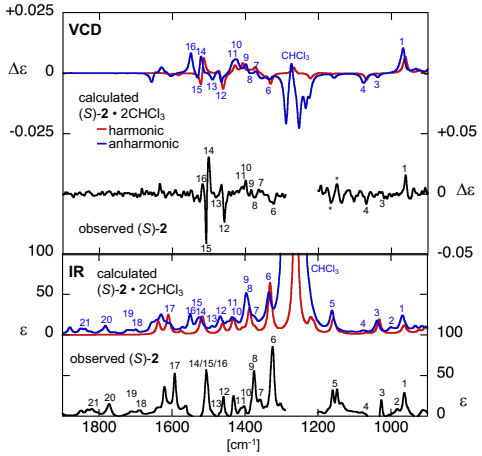


Figure S4: VCD (top) and IR (bottom) of measured (black) for (*S*)-**2** and calculated (red for harmonic DFT and blue for anharmonic DFT) spectra (*S*)-**2**•2$CHCl_3$. Part of corresponding VCD and IR peaks of observed spectra and anharmonically calculated spectra are labelled, while nonmatching peaks are indicated by asterisks. Both the anharmonic and harmonic VCD spectra were qualitatively similar to observed one, whereas only anharmonically calculated IR spectrum well reproduced the observed anharmonic signals in the region above 1600 cm^{-1} (peaks 18-21). Measurement conditions: *c* 0.3 M in $CHCl_3$, *l* 50 μm, corrected by solvent spectra. Calculation conditions: B3PW91/6-311++G(d,p). Harmonic calculated spectra are scaled by a factor of 0.98, while anharmonic ones are unscaled. Part of observed spectra are omitted due to strong solvent absorption.

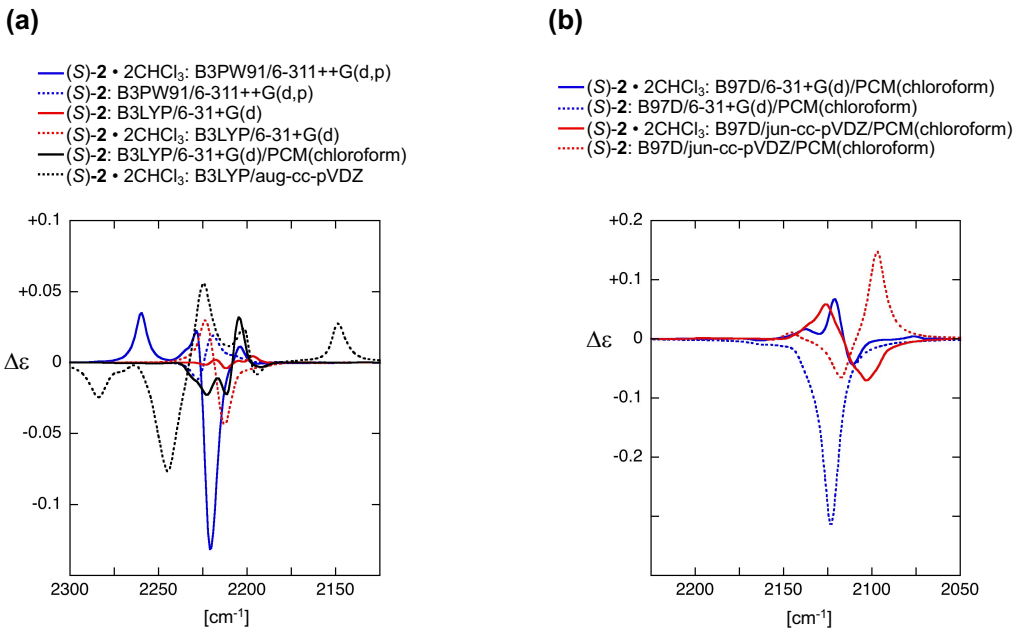


Figure S5: Predicted VCD patterns of (*S*)-**2** by anharmonic DFT using (a) B3PW91 or B3LYP, or (b) B97D functionals. All spectra were scaled by a factor of 0.967.

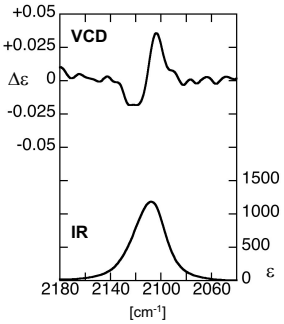
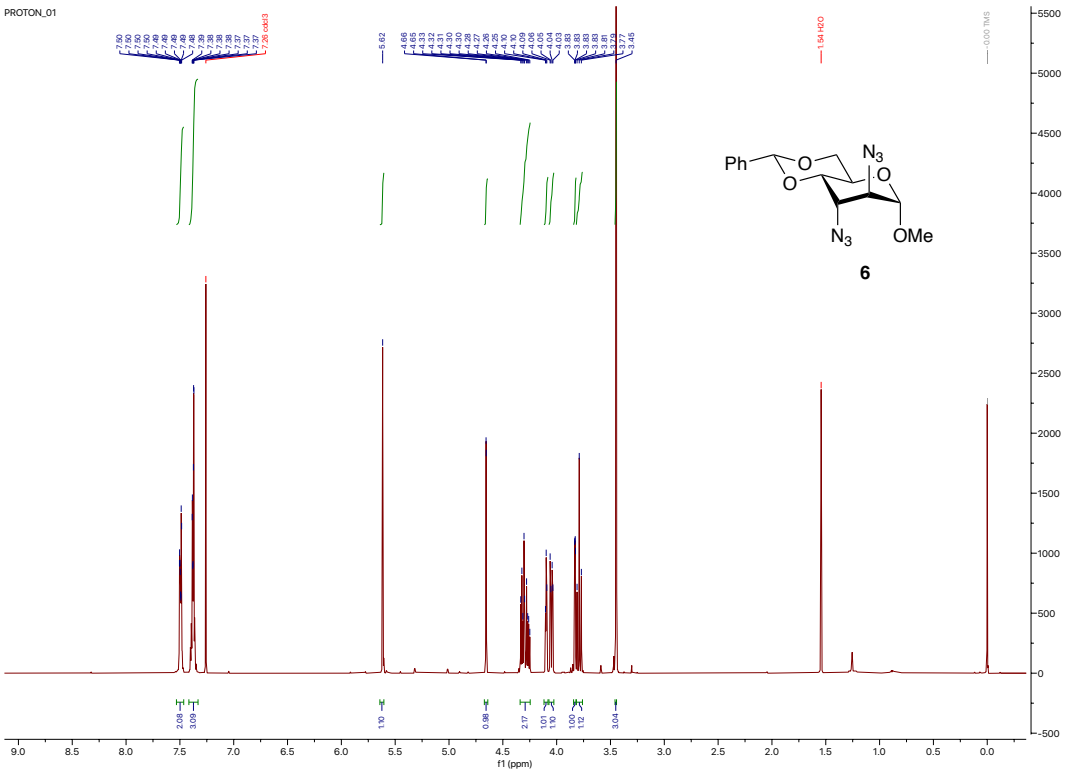
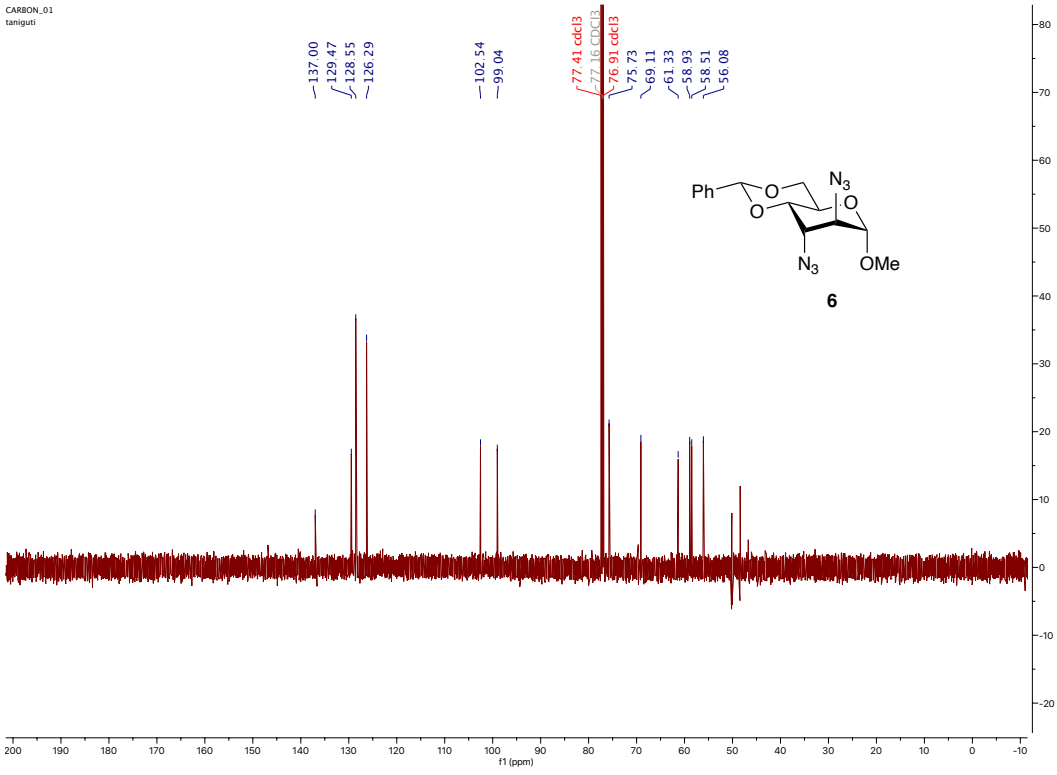


Figure S6: Observed VCD and IR spectra of 6 in DMSO-h_6 – H_2O (9:1) mixed solvent. Measurement conditions: *c* 0.04 M, *l* 50 μm, corrected by solvent spectra.

NMR Spectra


PROTON_01
Ph
O
O
N3
O
N3
OMe
6
f1 (ppm)


CARBON_01
taniguti
137.00
129.47
128.55
126.29
102.54
99.04
77.41 cdcl3
77.16 CDCl3
76.91 cdcl3
75.73
69.11
61.33
58.93
58.51
56.08
Ph
O
O
N3
O
N3
OMe
6
f1 (ppm)

2.8 References

1. N. Harada and K. Nakanishi, *J. Am. Chem. Soc.*, 1969, **91**, 3989-3991.
2. N. Harada and K. Nakanishi, *Circular dichroic spectroscopy : exciton coupling in organic stereochemistry*, University Science Books, 1983.
3. N. Harada, K. Nakanishi and N. Berova, in *Comprehensive Chiroptical Spectroscopy*, John Wiley & Sons, Inc., 2012, pp. 115-166.
4. N. Berova, L. D. Bari and G. Pescitelli, *Chem. Soc. Rev.*, 2007, **36**, 914-931.
5. N. Harada, *J. Am. Chem. Soc.*, 1973, **95**, 240-242.
6. T. F. Molinski, M. N. Salib, A. N. Pearce and B. R. Copp, *J. Nat. Prod.*, 2019, **82**, 1183-1189.
7. M. Satake, A. Morohashi, H. Oguri, T. Oishi, M. Hirama, N. Harada and T. Yasumoto, *J. Am. Chem. Soc.*, 1997, **119**, 11325-11326.
8. S. Abbate, F. Lebon, G. Longhi, S. E. Boiadjiev and D. A. Lightner, *J. Phys. Chem. B*, 2012, **116**, 5628-5636.
9. Y.-M. Shi, K. Hu, G. Pescitelli, M. Liu, X.-N. Li, X. Du, W.-L. Xiao, H.-D. Sun and P.-T. Puno, *Org. Lett.*, 2018, **20**, 1500-1504.
10. S. Matile, N. Berova, K. Nakanishi, S. Novkova, I. Philipova and B. Blagoev, *J. Am. Chem. Soc.*, 1995, **117**, 7021-7022.
11. K. Tsubaki, K. Takaishi, H. Tanaka, M. Miura and T. Kawabata, *Org. Lett.*, 2006, **8**, 2587-2590.
12. M. Balaz, J. D. Steinkruger, G. A. Ellestad and N. Berova, *Org. Lett.*, 2005, **7**, 5613-5616.
13. M. Balaz, A. E. Holmes, M. Benedetti, P. C. Rodriguez, N. Berova, K. Nakanishi and G. Proni, *J. Am. Chem. Soc.*, 2005, **127**, 4172-4173.
14. R. Wang, C. Geiger, L. Chen, B. Swanson and D. G. Whitten, *J. Am. Chem. Soc.*, 2000, **122**, 2399-2400.
15. J. Tabei, M. Shiotsuki, F. Sanda and T. Masuda, *Macromolecules*, 2005, **38**, 9448-9454.
16. G. Holzwarth and I. Chabay, *J. Chem. Phys.*, 1972, **57**, 1632-1635.
17. T. Taniguchi and K. Monde, *J. Am. Chem. Soc.*, 2012, **134**, 3695-3698.
18. T. Taniguchi, *Bull. Chem. Soc. Jpn.*, 2017, **90**, 1005-1016.
19. T. Hongen, T. Taniguchi, S. Nomura, J. Kadokawa and K. Monde, *Macromolecules*, 2014, **47**, 5313-5319.

20. K. Komori, T. Taniguchi, S. Mizutani, K. Monde, K. Kuramochi and K. Tsubaki, *Org. Lett.* 2014, **16**, 1386-1389.
21. G. Szilvagyi, B. Brem, G. Bati, L. Tolgyesi, M. Hollosi and E. Vass, *Dalton Trans.*, 2013, **42**, 13137-13144.
22. T. Asai, T. Taniguchi, T. Yamamoto, K. Monde and Y. Oshima, *Org. Lett.*, 2013, **15**, 4320-4323.
23. T. Asai, S. Morita, T. Taniguchi, K. Monde and Y. Oshima, *Org. Biomol. Chem.*, 2016, **14**, 646-651.
24. T. Taniguchi, D. Manai, M. Shibata, Y. Itabashi and K. Monde, *J. Am. Chem. Soc.*, 2015, **137**, 12191-12194.
25. C. I. Bautista-Hernández, R. E. Cordero-Rivera, E. A. Zúñiga-Estrada, N. Trejo-Carbajal, M. Meléndez-Rodríguez, O. R. Suárez-Castillo, M. Sánchez-Zavala, M. S. Morales-Ríos and P. Joseph-Nathan, *Tetrahedron Asymm.*, 2016, **27**, 623-638.
26. M. R. Poopari, Z. Dezhahang, K. Shen, L. Wang, T. L. Lowary and Y. Xu, *J. Org. Chem.*, 2015, **80**, 428-437.
27. Y. Hayashi, K. Nagai and S. Umemiya, *Eur. J. Org. Chem.*, 2019, **2019**, 678-681.
28. Y. Saito, M. Satake, R. Mori, M. Okayasu, H. Masu, M. Tominaga, K. Katagiri, K. Yamaguchi, S. Kikkawa, H. Hikawa and I. Azumaya, *Org. Biomol. Chem.*, 2020, **18**, 230-236.
29. K. Takaishi, K. Iwachido, R. Takehana, M. Uchiyama and T. Ema, *J. Am. Chem. Soc.*, 2019, **141**, 6185-6190.
30. T. Wu and X. You, *J. Phys. Chem. A*, 2012, **116**, 8959-8964.
31. C. L. Covington, V. P. Nicu and P. L. Polavarapu, *J. Phys. Chem. A*, 2015, **119**, 10589-10601.
32. V. P. Nicu, *Phys. Chem. Chem. Phys.*, 2016, **18**, 21202-21212.
33. S. Abbate, G. Mazzeo, S. Meneghini, G. Longhi, S. E. Boiadjiev and D. A. Lightner, *J. Phys. Chem. A*, 2015, **119**, 4261-4267.
34. S. Abbate, T. Bruhn, G. Pescitelli and G. Longhi, *J. Phys. Chem. A*, 2017, **121**, 394-400.
35. R. A. G. D. Silva, J. Kubelka, P. Bour, S. M. Decatur and T. A. Keiderling, *Proc. Nat. Acad. Sci.*, 2000, **97**, 8318-8323.
36. H.-Z. Tang, B. M. Novak, J. He and P. L. Polavarapu, *Angew. Chem. Int. Ed.*, 2005, **44**, 7298-7301.
37. R. Gangemi, G. Longhi, F. Lebon, S. Abbate and L. Laux, *Monatsh. Chem.*, 2005, **136**, 325-345.

38. L. Laux, V. Pultz, S. Abbate, H. A. Havel, J. Overend, A. Moscowitz and D. A. Lightner, *J. Am. Chem. Soc.*, 1982, **104**, 4276-4278.
39. A. T. Krummel and M. T. Zanni, *J. Phys. Chem. B*, 2008, **112**, 1336-1338.
40. K.-I. Oh, J.-H. Lee, C. Joo, H. Han and M. Cho, *J. Phys. Chem. B*, 2008, **112**, 10352-10357.
41. C. Fang, J. D. Bauman, K. Das, A. Remorino, E. Arnold and R. M. C. F. p. d. F. Hochstrasser, *Proc. Nat. Acad. Sci.*, 2008, **105**, 1472-1477.
42. U. Narayanan, T. A. Keiderling, C. J. Elsevier, P. Vermeer and W. Runge, *J. Am. Chem. Soc.*, 1988, **110**, 4133-4138.
43. U. Narayanan and T. A. Keiderling, *J. Am. Chem. Soc.*, 1988, **110**, 4139-4144.
44. T. Taniguchi, T. Suzuki, H. Satoh, Y. Shichibu, K. Konishi and K. Monde, *J. Am. Chem. Soc.*, 2018, **140**, 15577-15581.
45. M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, and D. J. Fox, *Gaussian 16*, *Rev A.03*, Wallingford CT, USA, 2016.
46. *Spartan 18*, Wavefunction Inc., Irvine, CA, USA, 2014.
47. G. Holzwarth, E. C. Hsu, H. S. Mosher, T. R. Faulkner and A. Moscowitz, *J. Am. Chem. Soc.*, 1974, **96**, 251-252.
48. P. L. Polavarapu, L. A. Nafie, S. A. Benner and T. H. Morton, *J. Am. Chem. Soc.*, 1981, **103**, 5349-5354.
49. P. Malon, L. J. Mickley, K. M. Sluis, C. N. Tam, T. A. Keiderling, S. Kamath, J. Uang and J. S. Chickos, *J. Phys. Chem.*, 1992, **96**, 10139-10149.
50. C. Shen, G. h. Loas, M. Srebro-Hooper, N. Vanthuyne, L. Toupet, O. Cador, F. Paul, J. T. López Navarrete, F. J. Ramírez, B. Nieto-Ortega, J. Casado, J.

Autschbach, M. Vallet and J. Crassous, *Angew. Chem. Int. Ed.*, 2016, **55**, 8062-8066.

51. A. J. Schmitz, D. G. Hogle, X. S. Gai, E. E. Fenlon, S. H. Brewer and M. J. Tucker, *J. Phys. Chem. B*, 2016, **120**, 9387-9394.
52. Enantiomerically pure (+)-**2a** was reported in only one paper. While this reference depicted the structure of (*R*)-**2a**, it was labelled as (*S*)-enantiomer. We thus unambigously determined the absolute configuration of (+)-**2a** as *R* by using VCD spectroscopy as shown in Figure S3. J. Liu, H.-X. Zheng, C.-Z. Yao, B.-F. Sun and Y.-B. Kang, *J. Am. Chem. Soc.*, 2016, **138**, 3294-3297.
53. C. Merten, J. Bloino, V. Barone and Y. Xu, *J. Phys. Chem. Lett.*, 2013, **4**, 3424-3428.
54. M. Fusè, G. Mazzeo, G. Longhi, S. Abbate, M. Masi, A. Evidente, C. Puzzarini and V. Barone, *J. Phys. Chem. B*, 2019, **123**, 9230-9237.
55. C. Merten, *Phys. Chem. Chem. Phys.*, 2017, **19**, 18803-18812.
56. J. Y. Park, H.-J. Kwon, S. Mondal, H. Han, K. Kwak and M. Cho, *Phys. Chem. Chem. Phys.*, 2020, **22**, 19223-19229.
57. N. J. Agard, J. A. Prescher and C. R. Bertozzi, *J. Am. Chem. Soc.*, 2004, **126**, 15046-15047.
58. T. Hoshi, E. Nozawa, M. Katano, T. Suzuki, H. Hagiwara, *Tetrahedron Lett.* 2004, **45,** 3485.
59. L. Zhao, D. Qi, K. Wang, T. Wang, B. Han, Z. Tang, J. Jiang, *Sci. Rep.* 2016, **6**, 28026.
60. M. Vondenhof, J. Mattay, *Chem. Ber.* 1990, **123,** 2457.
61. C. Bartolome, J. Garcia-Cuadrado, Z. Ramiro, P. Espinet, *Inorg. Chem.* 2010, **49**, 9758.
62. Y. Yamamoto, T. Hagiwara, H. Yamazaki, *Inorg. Chim. Acta* 1986, **115**, L35.
63. G. R. Schaller, F. Topić, K. Rissanen, Y. Okamoto, J. Shen, R. Herges, *Nature Chem.* 2014, **6**, 608.
64. Y. Takeda, M. Okazaki, S. Minakata, *Chem. Commun.* 2014, **50**, 10291.
65. W. Huang, Y. C. Zhang, R. Jin, B. L. Chen, Z. Chen, *Organometallics* 2018, **37,** 3196.
66. R. Aissaoui, A. Nourry, A. Coquel, T. T. H. Dao, A. Derdour, J.-J. Helesbeux, O. Duval, A.-S. Castanet, J. Mortier, *Org. Chem.* 2012, **77**, 718.
67. Y. L. Tnay, C. Chen, Y. Y. Chua, L. Zhang, S. Chiba, *Org. Lett.* 2012, **14**, 3550.
68. J. Liu, H.-K. Zheng, C.-Z. Yao, B.-F. Sun, Y.-B. Kang, *J. Am. Chem. Soc.* 2016,

138, 3294.

69. C.-T. Chen, S. -S. Weng, J. -Q. Kao, C. -C. Lin, M. -D. Jan, *Org. Lett.* 2005, **7,** 3343.

70. A. Franconetti, E. Álvarez, F. Cabrera-Escribano, *Synlett* 2017, **28**, 201.

CHAPTER 3: DEUTERIUM LABELLING TO EXTRACT LOCAL STEREOCHEMICAL INFORMATION BY VCD SPECTROSCOPY IN C-D STRETCHING REGION: A CASE STUDY OF SUGARS

3.1 Abstract

Stereochemical elucidation of molecules with multiple chiral centers is difficult. Even with VCD spectroscopy, excluding all but one diastereomeric structural candidates is challenging because stereochemical inversion of one chiral center among many centers does not always result in noticeable differences in their VCD spectra. This work demonstrates that introduction of a suitable VCD chromophore with absorption in the 2300-1900 cm^{-1} region can be used for extracting local stereochemical information and for stereochemical assignment of C-1 position of various sugars as a case study. Through studies on a series of epimeric pairs of monosaccharides and their derivatives, we found that the introduction of one -OCD_3 group to each C-1 position produced almost mirror-image VCD patterns in the 2300-1900 cm^{-1} region depending on the C-1 stereochemistry irrespective of the other molecular moiety. This work also shows that comparison of the observed VCD signals and calculated ones enables the stereochemical assignment of a chiral center in the vicinity of the chromophore. This study provides a proof of concept that the use of a VCD chromophore in the 2300-1900 cm^{-1} region enables the analysis of selected stereochemistry of suitable molecular systems. Further studies on this concept should lead to development of a method useful for the structural elucidation of other types of complex molecules.

3.2 Introduction

Middle-sized natural products (M.W. 500 to 2,000) such as oligosaccharides, cyclic peptides, macrolides, and other metabolites are promising drug candidates beyond Lipinski's rule of 5. Because middle-sized molecules target, for example, protein-protein interactions more efficiently than small molecules do, many such drugs have been approved and under development in recent years. [1, 2] For exploration of such natural products and their synthetic derivatives, their structural elucidation is one of the major bottlenecks as they possess multiple chiral centers. NMR coupling analysis and 2D NMR may establish the atom connectivity. However, such traditional analysis usually does not exclude all but one diastereomeric structural candidates. As a result, multiple analytical methods (e.g., Mosher-Kusumi method,[3] NMR DP4 analysis,[4] electronic circular dichroism spectroscopy,[5, 6] vibrational circular dichroism (VCD) spectroscopy,[6-9] Raman optical activity (ROA) spectroscopy,[7, 8] and chemical correlation) are applied to elucidate the stereochemistry of as many chiral centers in target molecules as possible.[10-12] Crystallization of middle-sized molecules suited for structural analysis by X-ray crystallography is not tedious. Crystalline sponge method[13] and cryoEM method microED[14] are emerging powerful analysis techniques for small organic molecules, but their applicability to middle-sized, flexible molecules are yet to be studied. New analytical methods that complement existing ones should facilitate studies involving middle-sized molecules.

VCD spectroscopy has become one of the most used methods to assign the absolute configuration of natural products and their synthetic derivatives.[6-9] Stereochemical elucidation of molecules with more than one chiral centre is possible, as diastereomers of small molecules often show satisfactorily different VCD patterns.[15-18] Since VCD signals originate from vibrational transitions, functional groups in the vicinity

of the chiral centre of interest tend to show characteristic signals that is indicative of its stereochemistry. For example, Merten and co-workers reported that some of anthrone and oxanthrone glucosides show characteristic VCD signals in the 1650-1500 cm^{-1} region that may serve as a marker for the configuration of the C-10 stereocenter.[19] Hereabout, Batista Jr. and coworkers suggested that spectral patterns at around 1390 cm^{-1} observed for acetonized 1,3-diols are markers for the absolute configuration of syn- and anti-oriented hydroxyl groups.[20] Such spectra-structures relationships provide useful criteria to assign diastereomeric structures of small molecules. However, spectral differences may be obscured for multi-functionalized middle-size molecules with many chiral centres, as reversal of a single chiral centre (an epimer) may not produce a noticeable spectral change. This point was also discussed by Reiher and co-workers in their studies on the limitation of ROA spectroscopy to distinguish molecules with many chiral centres.[21]

Common organic molecules do not show strong absorption in the 2300-1900 cm^{-1} region. Use of VCD spectroscopy in this region is a promising approach to extract stereochemical information from middle-sized molecules. In continuation of our studies on VCD couplet,[22-25] we recently showed that interaction of two azido groups produce a strong VCD couplet (at ca. 2100 cm^{-1}) whose shape depends on the local environment around the azido groups.[26] Although this finding should be useful in the analysis of, for example, the conformation of biomolecules, the necessity of introducing two azido groups to suitable positions limits its application to multi-functionalized middle-sized molecules. If one chromophore, not two, in the 2300-1900 cm^{-1} region is found sufficient to extract local stereochemical information and if the resultant VCD pattern is predictable by theoretical calculations, these should pave a way to develop new methods to elucidate the selected stereochemistry of target molecules. To prove this concept, one needs to prepare a series of diastereomeric pairs with a suitable VCD chromophore and confirm that their VCD spectra in the 2300-1900 cm^{-1} region exhibit drastically different, ideally

mirror-image, patterns irrespective of the other chiral centres. To provide a proof of concept, this work studied a methoxy-d_3 group (-OCD_3) installed at the C-1 position of a series of mono- and disaccharides and their derivatives. Sugars are chosen as a model system because of their diverse stereochemistry and functional groups, but we hope to apply the concept proven here to other types of molecules in future studies. Herein, we show the feasibility of the concept to determine the absolute configuration of the chiral centre, the C-1 stereochemistry of sugars in this case study, in the vicinity of an introduced chromophore by VCD spectroscopy in the 2300-1900 cm^{-1} region.

3.3 Result and discussion

3.3.1 Methyl-d_3 glycopyranosides

We started with studying monosaccharides possessing a -OCD_3 group. Considering the biological abundance and importance, D-glucose (**D-Glc**) and its enantiomer (**L-Glc**), D-galactose (**D-Gal**), D-xylose (**D-Xyl**), L-arabinose (**L-Ara**), D-mannose (**D-Man**), and L-fucose (**L-Fuc**) were selected. For each epimeric pair of methyl-d_3 glycopyranosides, their C-1 configuration is denoted as α or β (**Figure 1**). Most of the methyl-d_3 glycopyranosides were synthesized by Fischer glycosidation to produce α-β mixtures (free monosaccharides in CD_3OD stirred with catalytic acid), benzoylation (to facilitate diastereoseparation), chromatographic separation, and deprotection of benzoyl groups. Sterically and stereoelectronically unfavoured β-mannoside **CD_3-β-D-Man** was synthesized with using Crich β-mannosylation.[27]

Obtained methyl-d_3 glycopyranosides were dissolved in DMSO-h_6 and placed in a 50-μm BaF_2 cell for VCD measurement in the 2300-1900 cm^{-1} region. Note that DMSO-h_6, not DMSO-d_6, was used because the latter shows strong solvent absorption in this region due to C-D stretching vibrations. **Figure 1a** shows the VCD and IR spectra of **CD_3-α-D-Glc** and **CD_3-β-D-Glc**. For both epimers, two VCD bands were detected at 2065 cm^{-1} and around 2210 cm^{-1}. These signals were respectively assigned as symmetric (ν_s) and asymmetric (ν_{as}) CD_3 stretching vibrations.[28] To our surprise, these epimers showed almost mirror-image VCD patterns: **CD_3-α-D-Glc** showed a positive VCD signal for ν_s CD_3 and a negative one for ν_{as} CD_3, while **CD_3-β-D-Glc** showed a negative signal for ν_s CD_3 and a positive one for ν_{as} CD_3. It should be emphasized that such VCD patterns are exhibited despite the common presence of other four chiral centers (C2 to C5). Encouraged by this result, we studied the VCD of other monosaccharides. Changes in the configuration at C-4 did not affect the sign of the VCD signals of ν_s and ν_{as} CD_3 (**Figure 3.1b**).

Loss of the hydroxymethyl group at C-5 resulted in the reduction of the intensity of the n_s CD_3 signal for the 1*S* series (**CD_3-α-D-Man** and **CD_3-β-L-Ara**) but maintained the overall sign patterns (**Figure 3.1 c** and **d**). Reversal of the C-2 configuration also did not change the sign pattern for both 1*S* and 1*R* epimers (**Figure 3.1e**). Stereochemical inversions of C-3 and C-5 as well as the substitution of C-5 hydroxymethyl to methyl group were also tolerated (**Figure 3.1 f** and **g**). For all the tested monosaccharides in **Figure 3.1**, the 1*S* epimers showed a positive VCD signal for ν_s CD_3 and a negative one for ν_{as} CD_3, whereas the 1*R* epimers showed a negative signal for ν_s CD_3 and a positive one for ν_{as} CD_3. Thus, C-1 epimeric pairs exhibited mirror-image VCD patterns irrespective of the changes in the configurations and substituents at other positions. This indicated that extraction of the stereochemical information of a selected site from

molecules with multiple chiral centers is possible for suitable cases with the use of a suitable VCD chromophore in the 2300-1900 cm^{-1} region.

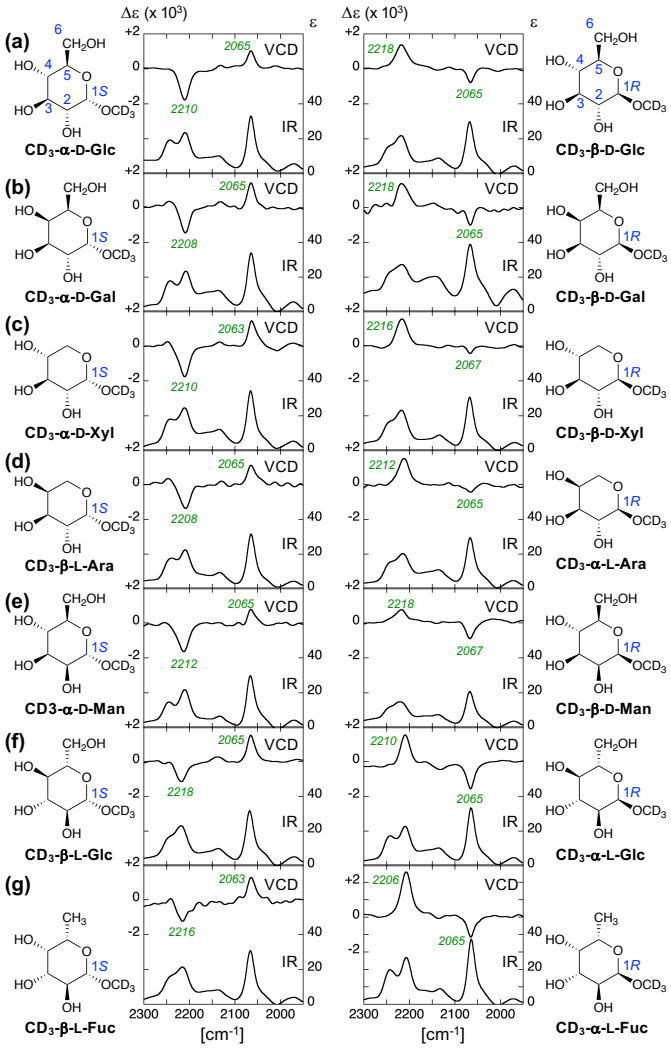


Figure 3.1: Structures and observed VCD and IR spectra of methyl-d_3 glycopyranosides derived from (a) D-glucose, (b) D-galactose, (c) D-xylose, (d) L-arabinose, (e) D-mannose, (f) L-glucose, and (g) L-fucose with (left) 1S configuration and (right) 1*R* configuration. Wavenumbers of VCD signal extrema for ν_s CD_3 (lower frequency) and ν_{as} CD_3 (higher frequency) are labelled in italic (green). Measurement conditions: *c* 0.5 – 2.0 M in DMSO-h_6, l50 µm.

3.3.2 Theoretical VCD calculations

Next, we tested the feasibility of determining the stereochemistry at C-1 solely by comparison of the observed CD_3 VCD signals and calculated signals in the 2300-1900 cm^{-1} region without prior knowledge of the NMR-determined C-1 stereochemistry. VCD spectra of representative **CD_3-α-D-Glc** and **CD_3-β-D-Glc** were calculated at the B3LYP/6-31G(d) level of theory with or without using polarizable continuum model (PCM) for DMSO. Both gas phase and PCM(DMSO) conditions predicted similar VCD spectra for each molecule (**Figure 3.2a**). Importantly, these spectra well reproduced the observed VCD patterns: **CD_3-α-D-Glc** showed a positive signal for ν_s CD_3 and a negative one for ν_{as} CD_3, whereas **CD_3-β-D-Glc** showed a negative signal for ν_s CD_3 and a positive one for ν_{as} CD_3. These results proved it feasible to determine the stereochemistry of a chiral center in the vicinity of a VCD chromophore in 2300-1900 cm^{-1} with using theoretical VCD calculations.

Chromophores in the 2300-1900 cm^{-1} region are advantageous for studying hydrophilic biomolecules. VCD measurements of H_2O solutions below 1900 cm^{-1} must be carried out in a special sample cell with a short pathlength (10 μm or shorter) and accordingly at very high sample concentration to circumvent its strong solvent absorption. With H_2O less absorbing in the 2300-1900 cm^{-1} region, VCD study in this region can be performed with a longer pathlength and a slightly lower concentration. **Figure 3.2b** shows the VCD spectra of H_2O solutions of **CD_3-α-** and **β-D-Glc** in a 25-μm CaF_2 cell. These VCD showed the same sign patterns to those observed for the DMSO-h_6 solutions. The observed VCD spectra in H_2O also were well reproduced by DFT calculations with implicit (PCM) as well as explicit (two-layer quantum mechanics (QM)/ molecular mechanics (MM) method for several MD snapshots with ca. 120 water molecules) models

for H_2O (**Figure 3.2a and Figure S1**). Note that the sample concentration required for reliably detecting CD_3 VCD signals is still much higher (3 M in this case) than that for ordinary VCD measurements below 1800 cm^{-1}.[15-18]

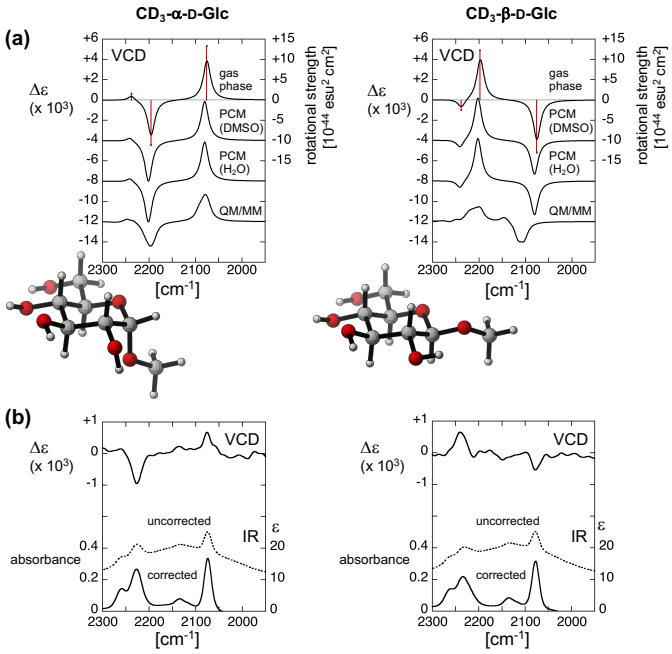


Figure 3.2: Calculated VCD spectra using several conditions and observed VCD/IR spectra for H_2O solution of **CD_3-α-D-Glc** (left) and **CD_3-β-D-Glc** (right). (a) VCD spectra predicted in the gas phase, with PCM for DMSO, with PCM for water, and with explicit water molecules. Red bars indicate the rotational strength of the predicted vibrational transitions for the most stable conformers in the gas phase. The most stable conformers predicted using PCM for water are drawn. DFT calculation conditions: B3LYP/6-31G(d) with or without PCM. See Supplementary Information for QM/MM calculation conditions. Frequency scaling factor: 0.955 for all the calculated spectra. (b) Observed VCD spectra corrected by a solvent VCD spectrum in De, uncorrected raw IR spectra in absorbance, and observed IR spectra corrected by a solvent IR spectrum in e. Corrected IR spectra in some wavenumber region showed negative values possibly due to solvent-sample interactions and are presented without further data processing. Measurement conditions: *c* 3.0 M in H_2O, *l* 25 μm.

While we initiated this work with the hope of seeing enough spectral differences for distinguishing epimers, such mirror-image patterns ideal for stereochemical assignment were far better than expected. In many reported cases, chromophores in a pair of epimers do not always produce mirror-image VCD signals due to influences from the

rest of the molecule.[15-18,20] Minor perturbation to the CD_3 VCD signals may be ascribed to the highly localized vibrational nature of the ν_s and ν_{as} CD_3 modes, as found for calculated **CD_3-α-** and **β-D-Glc** (**Figure S2**). Independence of these CD_3 vibrations were also bolstered by theoretical calculations of methyl-d_3 glycopyranosides without substituents at C2-5 positions (models (*S*)-**1** and (*R*)-**1**). Theoretical VCD spectra of (*S*)-**1** and (*R*)-**1** exhibited the same VCD sign patterns as those observed for 1*S* and 1*R* series of monosaccharides, respectively (**Figure 3.3**). Moreover, CD_3 VCD signals were less affected by conformational differences. Pyranose ring in sugars is known to take two main chair conformations $^4C_1$ and $^1C_4$. The -OCD_3 group is oriented either axial or equatorial depending on the ring conformation. Inspection of the VCD spectra in **Figure 3.2** and the preferred chair conformations of these monosaccharides (**Figure S3**) suggested that the CD_3 VCD sign patterns were consistent for both the axially- and equatorially-oriented -OCD_3 groups. The consistency of the VCD sign patterns was also found for calculated spectra of both the $^4C_1$ and $^1C_4$ conformers of (*S*)- and (*R*)-**1** (**Figure 3.3**). Last, use of less polar solvent also yielded the same sign patterns. See **Figure S4** for the VCD spectra of **CD_3-α-D-Bz$_4$G**, **CD_3-β-D-Bz$_4$G**, **CD_3-α-D-6LG**, and **CD_3-β-D-6LG** (vide infra) in $CHCl_3$. Other samples studied here could not be dissolved in apolar solvents. Overall, these results revealed the highly independent nature of the CD_3 VCD signals from other part of the molecule, which gave rise to almost mirror-image VCD patterns between epimers. Although whether mirror-image patterns are seen for other chromophores and other molecular systems is yet to be studied, these results suggested the importance of isolated vibrational modes for application of this concept for other molecules. It should also be noted that similarity between the VCD features of experimental **CD_3-α-** and **β-D-Glc** and those of computational (*S*)- and (*R*)-**1** indicates the possibility of using an extremely simplified structure for calculations.

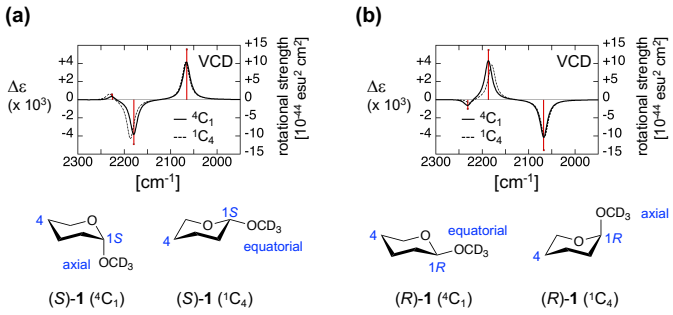


Figure 3.3: Calculated VCD spectra of simplified models (a) (*S*)-**1** and (b) (*R*)-**1** in $^4C_1$ and $^1C_4$ conformations. Red bars indicate the rotational strength of the predicted vibrational transitions for the $^4C_1$ conformers. Calculation conditions: DFT/B3LYP/6-31G(d). Frequency scaling factor: 0.955 for all the spectra.

3.3.3 Applicability of the concept on other complex sugar system

To showcase the wide applicability of the concept of extraction of local stereochemical information by -OCD_3 group, we studied the VCD spectra of the C-1 epimeric pairs of D-gentiobiose (β1,6-linked glucose dimer, **D-Gen**), 5-membered D-glucofuranoside (**D-*f*Glc**), and 2,3,4,6-*O*-tetrabenzoylated D-glucopyranoside (**D-Bz$_4$G**) (**Figure 3.4a**-c). **CD$_3$-α-D-Gen** was synthesized starting from **CD$_3$-α-D-Glc** with using Kahne sulfoxide glycosylation,[29] whereas **CD$_3$-β-D-Gen** was prepared from D-gentiobiose by a 3-step synthesis. Meanwhile, **CD$_3$-α-** and **β-D-*f*Glc** were obtained by $FeCl_3$-mediated glycosidation of D-glucose.[30] **D-Bz$_4$G** epimers are synthetic intermediates of **CD$_3$-α-** and **β-D-Glc**. Their VCD spectra in DMSO-h_6 are shown in Fig. 4a-c. All the α epimers (1*S* series) showed a positive ν_s CD_3 signal and a negative ν_{as} CD_3 one, whereas β (1*R*) counterparts showed the opposite pattern. Thus, their OCD_3 VCD signals well reflected the stereochemistry of interest even for a disaccharide with 10 chiral centers, a sugar with a different ring size, and a multi-chromophoric larger molecule. These results highlighted the potential of the concept of stereochemical assignment of the selected site for larger, complex molecular systems.

Isotopically labelled molecules are processed by enzymes often equally as their non-labelled counterparts. This permits biochemical utility of deuterium-containing chromophores for sample preparation and for tracking the stereochemical outcome of reactions. Such applications were demonstrated in a study on an epimeric pair of 6-*O*-lauroylated D-glucose (**6LG**) (**Figure 4d**). Non-labelled **CH$_3$-α-** and **β-6LG** are surfactant molecules with antimicrobial activity[31, 32] and have been prepared by lipase-catalyzed reactions from the corresponding **CH$_3$-α-** and **β-D-Glc**.[33] Following the procedure for the enzymatic conversion of methyl-h_3 species, we prepared **CD$_3$-α-** and **β-D-6LG** from **CD$_3$-α-** and **β-D-Glc**, respectively. The obtained lauroylated epimers showed VCD patterns well reflecting their C-1 stereochemistry (**Figure 4d**). From the viewpoint of stereochemical tracking, VCD spectroscopy confirmed that the C-1 stereochemistry was unaffected before and after the enzymatic transformation. With the insight obtained in this work, applications of VCD spectroscopy in the 2300-1900 cm^{-1} region for monitoring a stereochemical outcome of enzymatic as well as chemical reactions for other classes of molecules should be possible in future. One of such applications is currently under study by our group.

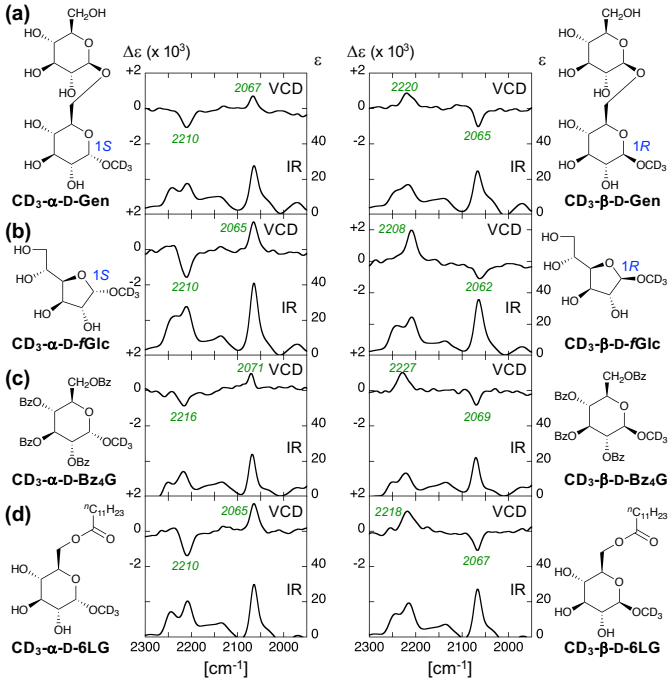


Figure 3.4: Structures and observed VCD and IR spectra of methyl-d_3 glycosides derived from (a) D-gentiobiose, (b) D-glucofuranose, (c) 2,3,4,6-O-tetrabenzoyl-D-glucose, and (d) 6-O-lauroyl-D-glucose with (left) 1*S* configuration and (right) 1*R* configuration. Wavenumbers of VCD signal extrema for ν_s CD_3 (lower frequency) and ν_{as} CD_3 (higher frequency) are labelled in italic (green). IR spectra of 6-O-lauroyl-D-glucose in some wavenumber region showed negative values possibly due to solvent-sample interactions and are presented without further data processing. Measurement conditions: *c* 0.5 or 1.0 M in DMSO-h_6, *l* 50 μm.

3.4 Conclusion

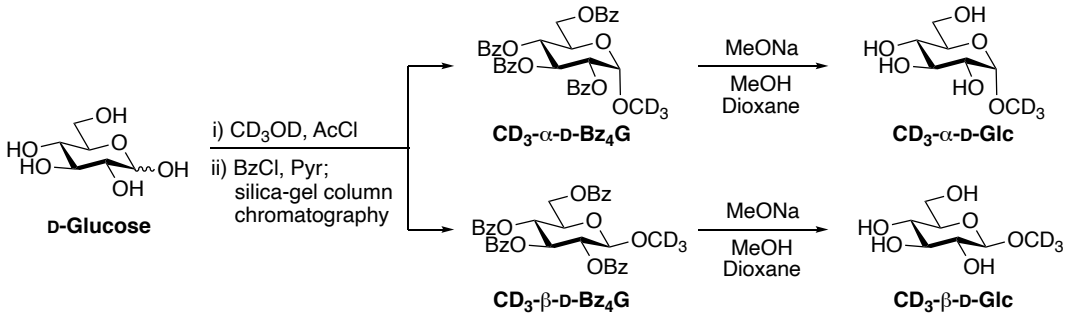
Stereochemical elucidation of multi-functionalized middle-sized molecules is difficult even with the use of several analysis methods. Through VCD studies on a series of methyl-d_3 glycoside, this work demonstrated that introduction of one VCD chromophore in the 2300-1900 cm^{-1} region enables extraction of local stereochemical information and determination of the stereochemistry of the chiral center of interest. Deuterium-containing VCD chromophores should be versatile when biochemical studies are involved. Although this work provided a proof of concept through studies on sugar molecules and -OCD_3 group, further studies on this concept should lead to development of a method useful for the structural elucidation of other types of complex molecules.

3.5 Experimental section

3.5.1 General experimental details

^{1}H NMR (500 MHz) and ^{13}C NMR (125 MHz) spectra were recorded on a Varian Inova instrument at 25 °C. Chemical shift values (δ) are reported in ppm relative to tetramethylsilane, $CDCl_3$ (^{13}C, δ 77.00), CD_3OD (^{1}H, δ 3.31), or D_2O (^{1}H, δ 4.75). The following abbreviations were used for signal multiplicities: s = singlet; d = doublet; t = triplet; and m = multiplet. Electrospray ionization mass spectrometry was measured on an Exactive Plus (Thermo Scientific). For vibrational spectroscopy, each sample was dissolved in $CHCl_3$, DMSO-h_6 or H_2O and placed in a 25-μm CaF_2 cell or a 50-μm BaF_2 cell. VCD and IR spectra were recorded using a JASCO FVS-6000 spectrometer with 8 cm^{-1} resolution for 1500 and 16 scans, respectively. To detect VCD signals in the 2300-1900 cm^{-1} region with higher S/N, the spectrometer was inserted an optical filter that passes through 2400-1900 cm^{-1} light and the signal was detected with an InSb detector.[26] The modulation frequency of the photoelastic modulator was set to 2127 cm^{-1}. All the VCD and IR spectra except the raw spectra in **Figure 3.2b** were corrected by solvent spectra obtained under the identical measurement conditions.

3.5.2 Synthesis of CD_3-α-D-Glc, CD_3-β-D-Glc, CD_3-α-D-Bz_4G, and CD_3-β-D-Bz_4G and General Procedures



(A) General procedure for Fischer glycosidation

CD_3OD (1 mL) was added AcCl (10 μL) and D-glucose (200 mg, 1.11 mmol) and the mixture was stirred overnight. After removal of the solvent, the crude mixture was used for the next reaction without purification unless otherwise noted.

(B) General procedure for benzoylation and the following chromatographic separation

A crude mixture obtained by the procedure (A) was added pyridine (5 mL) and benzoyl chloride (1 mL) and stirred overnight at 70 °C. The mixture was diluted with EtOAc, washed sequentially with 2M HCl aq, 2M NaOH aq, and brine, and then dried over $MgSO_4$. After removal of the solvent, the mixture was purified by silica-gel column chromatography (hexane-EtOAc = 3:1), which afforded **CD_3-α-D-Bz_4G** (302 mg, 45% in 2 steps) and **CD_3-β-D-Bz_4G** (130 mg, 19% in 2 steps). The ^{1}H NMR spectra of these compounds were virtually the same as those reported for methyl-h_3 2,3,4,6-tetra-*O*-benzoyl-α-D-glucopyranoside and methyl-h_3 2,3,4,6-tetra-*O*-benzoyl-β-D-glucopyranoside except for the absence of the -OCH_3 signal.[34]

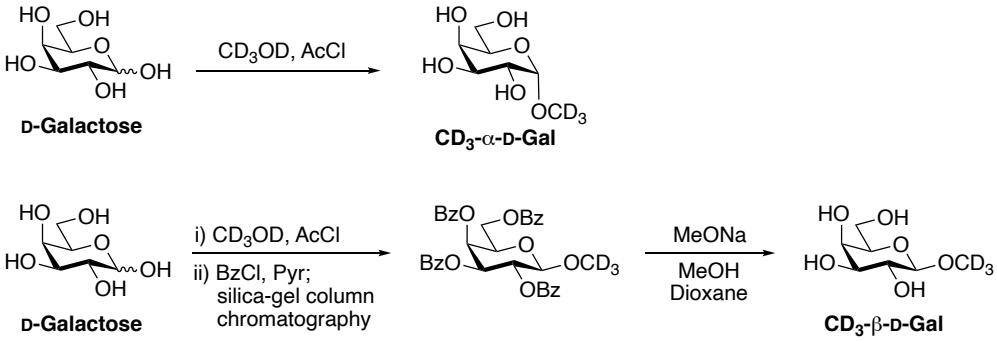
CD_3-α-D-Bz_4G: ^{1}H NMR ($CDCl_3$) δ 8.05 (d, ArH, *J* = 7.2 Hz, 2H), 7.99 (d, ArH, *J* = 7.2 Hz, 2H), 7.94 (d, ArH, *J* = 7.2 Hz, 2H), 7.87 (d, ArH, *J* = 7.2 Hz, 2H), 7.58-7.47 (m, ArH, 3H), 7.45-7.32 (m, ArH, 7H), 7.32-7.25 (m, ArH, 2H), 6.19 (dd, H-3, *J* = 10.0, 10.0 Hz, 1H), 5.69 (dd, H-4, *J* = 9.8, 9.8 Hz, 1H), 5.31 (dd, H-2, *J* = 3.4, 10.2 Hz, 1H), 5.26 (d, H-1, *J* = 3.7 Hz, 1H), 4.62 (dd, H-6a, *J* = 2.8, 12.1 Hz, 1H), 4.49 (dd, H-6b, *J* = 5.3, 12.2 Hz, 1H), 4.43 (m, H-5, 1H); **CD_3-β-D-Bz_4G**: ^{1}H NMR ($CDCl_3$) δ 8.02 (m, ArH, 2H), 7.97 (m, ArH, 2H), 7.90 (m, ArH, 2H), 7.82 (m, ArH, 2H), 7.57-7.47 (m, ArH, 3H), 7.44-7.32 (m, ArH, 7H), 7.30-7.26 (m, ArH, 2H), 5.91 (dd, H-3, *J* = 9.6, 9.6 Hz, 1H), 5.68 (dd, H-4, *J* = 9.8, 9.8 Hz, 1H), 5.52 (dd, H-2, *J* = 7.9, 9.8 Hz, 1H), 4.77 (d, H-1, *J* = 7.9 Hz, 1H), 4.65 (dd, H-6a, *J* = 3.2, 12.1 Hz, 1H), 4.51 (dd, H-6b, *J* = 5.4, 12.1 Hz, 1H), 4.16 (m, H-5, 1H).

(C) General procedure for deprotection of benzoyl or acetyl groups

CD_3-β-D-Bz_4G (125 mg, 204 μmol) in MeOH (2 mL) and 1,4-dioxane (2 mL) was added MeONa (20 μL) and stirred overnight at room temperature. The mixture was neutralized using Amberlite® 120IR (hydrogen form), then filtered and evaporated. The crude residue was added H_2O and Et_2O and the aqueous layer was collected and evaporated, which afforded **CD_3-β-D-Glc** (34 mg, 85%).[28]

CD_3-α-D-Glc[28] was prepared in a similar manner starting from **CD_3-α-D-Bz_4G**.

3.5.3 Synthesis of CD_3-α-D-Gal and CD_3-β-D-Gal

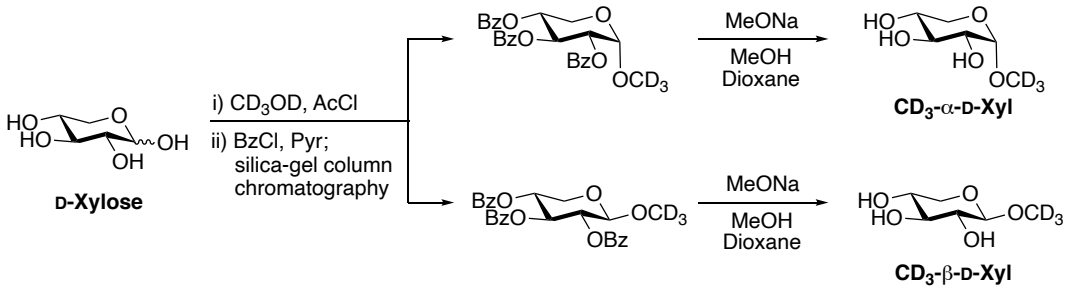


CD_3-α-D-Gal was obtained starting from D-galactose using the general procedure (A) and the following silica-gel column chromatography separation (chloroform-MeOH = 5:1). Its ^{1}H NMR spectrum was virtually the same as that observed for commercially available methyl-h_3 α-D-galactopyranoside except for the absence of the -OCH_3 signal. **CD_3-α-D-Gal**: ^{1}H NMR (D_2O) δ 4.81 (d, H-1, J = 2.4 Hz, 1H), 3.94 (dd, J = 1.3, 2.5 Hz, H-4, 1H), 3.87 (m, H-5, 1H), 3.83-3.76 (m, H-2, H-3, 2H), 3.74-3.69 (m, H-6a, H-6b, 2H).

CD_3-β-D-Gal was obtained starting from D-galactose using the general procedures (A)-(C). The ^{1}H NMR spectra of methyl-d_3 2,3,4,6-tetra-*O*-benzoyl-β-D-galactopyranoside and **CD_3-β-D-Gal** were virtually the same as those reported for methyl-h_3 2,3,4,6-tetra-*O*-benzoyl-β-D-galactopyranoside[34,35] and observed for commercially available methyl-h_3 β-D-galactopyranoside, respectively, except for the absence of the -OCH_3 signal. Methyl-d_3 2,3,4,6-tetra-*O*-benzoyl-β-D-galactopyranoside: ^{1}H NMR ($CDCl_3$) δ 8.09 (m, ArH, 2H), 8.03 (m, ArH, 2H), 7.97 (m, ArH, 2H), 7.78 (m, ArH, 2H), 7.65-7.54 (m, ArH, 2H), 7.53-7.36 (m, ArH, 8H), 7.25-7.21 (m, ArH, 2H), 6.00 (dd, H-4, J = 1.2, 3.5 Hz, 1H), 5.79 (dd, J = 7.9, 10.4 Hz, H-2, 1H), 5.61 (dd, H-3, J = 3.4, 10.3 Hz, 1H), 4.75 (d, H-1, J = 7.8 Hz, 1H), 4.70 (dd, H-6a, J = 6.6, 11.3 Hz, 1H), 4.43 (dd, H-6b, J = 6.7, 11.4 Hz, 1H), 4.33 (m, H-5, 1H). **CD_3-β-D-Gal**: ^{1}H NMR (CD_3OD) δ 4.13 (d, H-1, J = 7.1 Hz,

1H), 3.83 (d, H-4, J = 2.1 Hz, 1H), 3.79-3.70 (m, H-6a, H-6b, 2H), 3.53-3.43 (m, H-2, H-3, H-5, 3H).

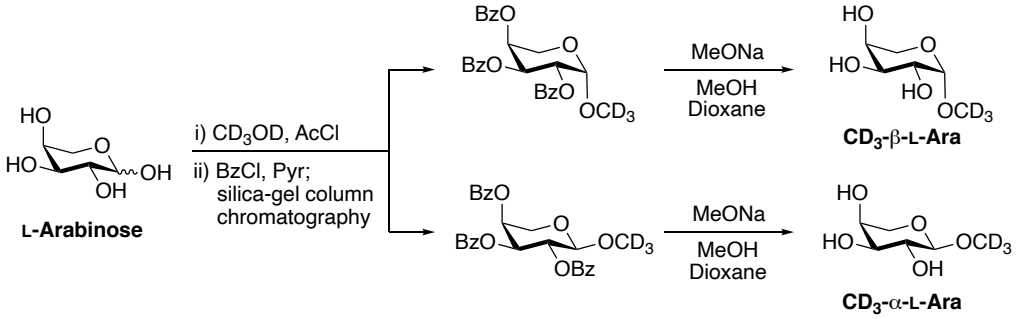
3.5.4 Synthesis of CD_3-α-D-Xyl and CD_3-β-D-Xyl



CD_3-α-D-Xyl and **CD_3-β-D-Xyl** were obtained starting from D-xylose using the general procedures (A)-(C). The ^{1}H NMR spectra of tribenzoyl synthetic intermediates were virtually the same as those reported for methyl-h_3 2,3,4-tri-*O*-benzoyl-α- and β-D-xylopyranoside except for the absence of the -OCH_3 signal,[36] while those of **CD_3-α-D-Xyl** and **CD_3-β-D-Xyl** were virtually the same as those observed for commercially available methyl-h_3 α- and β-D-xylopyranoside except for the absence of the -OCH_3 signal. Methyl-d_3 2,3,4-tri-*O*-benzoyl-α-D-xylopyranoside: ^{1}H NMR ($CDCl_3$) δ 7.98 (m, ArH, 4H), 7.92 (m, ArH, 2H), 7.51 (m, ArH, 2H), 7.44 (m, ArH, 1H), 7.38 (m, ArH, 4H), 7.31 (m, ArH, 2H), 6.16 (dd, H-3, J = 9.8, 9.8 Hz, 1H), 5.42 (m, H-4, 1H), 5.26 (dd, H-2, J = 3.4, 10.1 Hz, 1H), 5.17 (d, H-1, J = 3.5 Hz, 1H), 4.10 (dd, H-5a, J = 6.0, 10.8 Hz, 1H), 3.84 (dd, H-5b, J = 10.7, 10.7 Hz, 1H). Methyl-d_3 2,3,4-tri-*O*-benzoyl-β-D-xylopyranoside: ^{1}H NMR ($CDCl_3$) δ 8.02-7.94 (m, ArH, 6H), 7.55-7.45 (m, ArH, 3H), 7.40-7.31 (m, ArH, 6H), 5.80 (dd, H-3, J = 7.6, 7.6 Hz, 1H), 5.40 (dd, H-2, J = 5.7, 7.6 Hz, 1H), 5.33 (m, H-4, 1H), 4.74 (d, H-1, J = 5.7 Hz, 1H), 4.44 (dd, H-5a, J = 4.4, 12.0 Hz, 1H), 3.71 (dd, H-5b, J = 7.4, 12.2 Hz, 1H).**CD_3-α-D-Xyl**: ^{1}H NMR (CD_3OD) δ 4.61

(d, H-1, *J* = 3.6 Hz, 1H), 3.58-3.51 (m, H-3, H-4, 2H), 3.49-3.41 (m, H-5a, H-5b, 2H), 3.36 (dd, H-2, *J* = 3.6, 9.5 Hz, 1H). **CD_3-β-D-Xyl**: ^{1}H NMR (D_2O) δ 4.29 (d, H-1, *J* = 7.8 Hz, 1H), 3.93 (dd, H-5a, *J* = 5.2, 11.7, 1H), 3.58 (m, H-4, 1H), 3.40 (dd, H-3, *J* = 9.3, 9.3 Hz, 1H), 3.29 (dd, H-5b, *J* = 11.0, 11.0 Hz, 1H), 3.21 (dd, H-2, *J* = 7.8, 9.1 Hz, 1H).

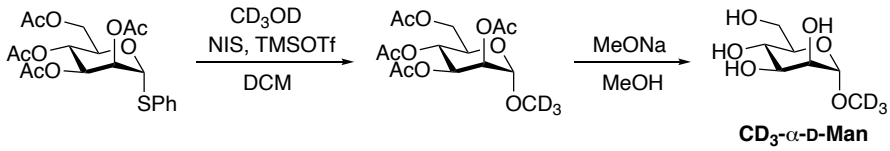
3.5.5 Synthesis of CD_3-β-L-Ara and CD_3-α-L-Ara



CD_3-β-L-Ara and **CD_3-α-L-Ara** were obtained starting from L-arabinose using the general procedures (A)-(C). The ^{1}H NMR spectra of tribenzoyl synthetic intermediates and **CD_3-β-L-Ara** were virtually the same as those reported for methyl-h_3 2,3,4-tri-*O*-benzoyl-β- and α-L-arabinopyranoside[35] and methyl-h_3 β-L-arabinopyranoside,[37] respectively, except for the absence of the -OCH_3 signal. The ^{1}H NMR spectrum of **CD_3-α-L-Ara** was virtually the same as that observed for commercially available methyl-h_3 α-L-arabinopyranosid except for the absence of the -OCH_3 signal. Methyl-d_3 2,3,4-tri-*O*-benzoyl-β-L-arabinopyranoside: ^{1}H NMR ($CDCl_3$) δ 8.11 (d, ArH, *J* = 7.3 Hz, 2H), 8.00 (d, ArH, *J* = 7.3 Hz, 2H), 7.85 (d, ArH, *J* = 7.3 Hz, 2H), 7.59 (t, ArH, *J* = 7.4 Hz, 1H), 7.54-7.40 (m, ArH, 5H), 7.37 (t, ArH, *J* = 7.7 Hz, 1H), 7.26 (t, ArH, *J* = 7.9 Hz, 2H), 5.96 (dd, H-3, *J* = 3.5, 10.7 Hz, 1H), 5.78-5.71 (m, H-2, H-4, 2H), 5.25 (d, H-1, *J* = 3.5 Hz, 1H), 4.18 (m, H-5a, 1H), 4.12 (dd, H-5b, *J* = 1.6, 13.3 Hz, 1H). Methyl-d_3 2,3,4-tri-*O*-benzoyl-α-L-arabinopyranoside: ^{1}H NMR ($CDCl_3$) δ 8.09-7.99 (m, ArH, 4H), 7.94-7.88

(d, ArH, J = 7.5 Hz, 2H), 7.60-7.51 (m, ArH, 2H), 7.50-7.38 (m, ArH, 5H), 7.35-7.29 (t, ArH, J = 7.8 Hz, 2H), 5.76-5.68 (m, H-2, H-4, 2H), 5.61 (dd, H-3, J = 3.4, 8.8 Hz, 1H), 4.67 (dd, H-1, J = 6.2 Hz, 1H), 4.33 (dd, H-5a, J = 3.7, 13.1 Hz, 1H), 3.91 (dd, H-5b, J = 1.6, 12.8 Hz, 1H). **CD_3-β-L-Ara**: ^{1}H NMR (D_2O) δ 4.79 (d, H-1, J = 2.4 Hz, 1H), 3.96 (m, H-4, 1H), 3.83 (d, H-5a, J = 11.6 Hz, 1H), 3.81-3.78 (m, H-2, H-3, 2H), 3.61 (dd, H-5b, J = 1.8, 12.7 Hz, 1H). **CD_3-α-L-Ara**: ^{1}H NMR (D_2O) δ 4.23 (d, H-1, J = 7.6 Hz, 1H), 3.93-3.86 (m, H-4, H-5a, 2H), 3.66-3.61 (m, H-3, H-5b, 2H), 3.50 (m, H-2, 1H).

3.5.6 Synthesis of CD_3-α-D-Man

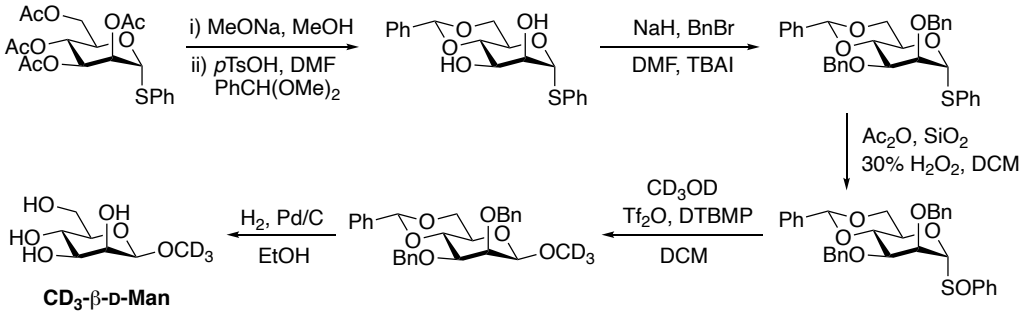


A solution of phenyl 2,3,4,6-tetra-*O*-acetyl-1-thio-α-D-mannopyranoside (570 mg, 1.30 mmol) in CH_2Cl_2 (5 mL) was added CD_3OD (160 μL) and stirred at 0 °C. To this mixture was added a solution of *N*-iodosuccinimide (750 mg) in THF (2.5 mL) dropwise, trimethylsilyl trifluoromethanesulfonate (25 μL) dropwise and stirred at 0 °C for 40 mins. The mixture was added triethylamine (20 μL), diluted with EtOAc, washed sequentially with a 1:1 mixed solution of 10% $Na_2S_2O_3$ aq and sat $NaHCO_3$ aq and then brine, and dried over $MgSO_4$. After removal of the solvent, the mixture was purified by silica-gel column chromatography (hexane-EtOAc = 5:2 to 1:1), which provided methyl-d_3 2,3,4,6-tetra-*O*-acetyl-α-D-mannopyranoside (122 mg, 26%). Its ^{1}H NMR spectrum was virtually the same as that observed for commercially available methyl-h_3 2,3,4,6-tetra-*O*-acetyl-α-D-mannopyranoside except for the absence of the -OCH_3 signal. Methyl-d_3 2,3,4,6-tetra-*O*-acetyl-α-D-mannopyranoside: ^{1}H NMR ($CDCl_3$) δ 5.34 (dd, H-3, J = 3.6, 9.8 Hz, 1H),

5.28 (dd, H-4, J = 9.8, 9.8 Hz, 1H), 5.24 (dd, H-2, J = 1.7, 3.2 Hz, 1H), 4.72 (d, H-1, J = 1.5 Hz, 1H), 4.29 (dd, H-6a, J = 5.4, 12.2 Hz, 1H), 4.13 (dd, H-6b, J = 2.4, 12.2 Hz, 1H), 3.97 (ddd, H-5, J = 2.4, 5.4, 9.8 Hz, 1H), 2.16 (s, Ac, 3H), 2.11 (s, Ac, 3H), 2.04 (s, Ac, 3H), 1.99 (s, Ac, 3H).

In a similar manner to the general procedure (C) without the use of 1,4-dioxane, methyl-d_3 2,3,4,6-tetra-*O*-acetyl-α-D-mannopyranoside (104 mg, 285 μmol) was converted to **CD_3-α-D-Man** (31 mg, 55%). Its ^{1}H NMR spectrum was virtually the same as that observed for commercially available methyl-h_3 α-D-mannopyranoside except for the absence of the -OCH_3 signal. **CD_3-α-D-Man**: ^{1}H NMR (D_2O) δ 4.74 (d, H-1, J = 1.5 Hz, 1H), 3.90 (dd, H-2, J = 2.0, 3.4 Hz, 1H), 3.87 (dd, H-6a, J = 2.0, 12.2 Hz, 1H), 3.75-3.71 (m, H-3, H-6b, 2H), 3.64-3.56 (m, H-4, H-5, 2H).

3.5.7 Synthesis of CD_3-β-D-Man



In a similar manner to the general procedure (C) without the use of 1,4-dioxane, phenyl 2,3,4,6-tetra-*O*-acetyl-1-thio-α-D-mannopyranoside (2.20 g, 5.00 mmol) was deacetylated. A solution of the crude compound in DMF (25 mL) was added benzaldehyde dimethoxy acetal (1 mL) and a catalytic amount of *p*-TsOH•H_2O and stirred overnight at room temperature. The mixture was diluted with $CHCl_3$, washed sequentially with sat $NaHCO_3$ aq. and brine, and dried over $MgSO_4$. After removal of the solvent, the

mixture was purified by recrystallization using EtOAc, which afforded phenyl 4,6-*O*-benzylidine-1-thio-α-D-mannopyranoside[38] (1.36 g, 67%).

A solution of phenyl 4,6-*O*-benzylidene-1-thio-α-D-mannopyranoside (720 mg, 2.00 mmol) in DMF (25 mL) was added 60% NaH in mineral oil (200 mg) and stirred at room temperature. After 15 mins, benzyl bromide (0.7 mL) was added dropwise followed by the addition of a catalytic amount of tetra-*n*-butylammonium iodide and stirred overnight at room temperature. The mixture was quenched with H_2O and was diluted with EtOAc. The organic phase was washed with brine and dried over $MgSO_4$. After removal of the solvent, the mixture was purified by silica-gel column chromatography (hexane-EtOAc = 25:1 to 15:1), which afforded phenyl 2,3-di-*O*-benzyl-4,6-*O*-benzylidine-1-thio-α-D-mannopyranoside[39] (293 mg, 27%).

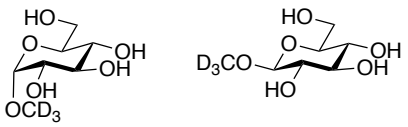
A solution of phenyl 2,3-di-*O*-benzyl-4,6-*O*-benzylidine-1-thio-α-D-mannopyranoside (270 mg, 500 μmol) in CH_2Cl_2 (5 mL) was added SiO_2 (100 mg), acetic anhydride (50 μL), and 30% H_2O_2 aq (110 μL) and stirred overnight at room temperature. The mixture was filtered through cotton, diluted with EtOAc, washed sequentially with 10% $Na_2S_2O_3$ aq, sat $NaHCO_3$ aq and brine, and dried over $MgSO_4$. After removal of the solvent, the residue was washed with methanol and hexane to yield phenyl 2,3-di-*O*-benzyl-4,6-*O*-benzylidine-1-thio-α-D-mannopyranoside *S*-oxide[40] (128 mg, 46%).

A solution of phenyl 2,3-di-*O*-benzyl-4,6-*O*-benzylidine-1-thio-α-D-mannopyranoside *S*-oxide (116 mg, 209 μmol) in CH_2Cl_2 (5 mL) was added 2,6-di-*tert*-butyl-4-methylpyridine (86 mg), trifluoromethanesulfonic anhydride (34 μL), CD_3OD (68 μL) and stirred at -78 °C. After 30 mins the reaction was quenched with the addition of Et_3N, diluted with EtOAc, washed sequentially with sat NH_4Cl aq, sat $NaHCO_3$ aq, and brine and dried over $MgSO_4$. After removal of the solvent, the mixture was purified by silica-gel column chromatography (hexane-EtOAc = 10:1 to 5:2), which provided methyl-d_3 2,3-di-*O*-benzyl-4,6-*O*-benzylidine-β-D-mannopyranoside (51 mg, 52%). Its ^1H NMR

spectrum was virtually the same as that reported for methyl-h_3 2,3-di-*O*-benzyl-4,6-*O*-benzylidine-β-D-mannopyranoside except for the absence of the -OCH_3 signal.[27] ^{1}H NMR ($CDCl_3$) δ 7.53-7.44 (m, ArH, 4H), 7.40-7.23 (m, ArH, 11H), 5.62 (s, CH, 1H), 4.96 (d, CH_2, *J* = 12.2 Hz, 1H), 4.84 (d, CH_2, *J* = 12.2 Hz, 1H), 4.68 (d, CH_2, *J* = 12.7, 1H), 4.58 (d, CH_2, *J* = 12.7, 1H), 4.47 (s, H-1, 1H), 4.32 (dd, H-6a, *J* = 4.9, 10.7 Hz, 1H), 4.20 (dd, H-4, *J* = 9.3, 9.8 Hz, 1H), 3.97-3.90 (m, H-2, H-6b, 2H), 3.58 (dd, H-3, *J* = 3.2, 10.0 Hz, 1H), 3.33 (ddd, H-5, *J* = 4.6, 9.6, 9.6 Hz, 1H).

A solution of methyl 2,3-di-*O*-benzyl-4,6-*O*-benzylidine-β-D-mannopyranoside (49 mg, 105 μmol) in EtOH (20 mL) was added a catalytic amount of 10% Pd/C and stirred at room temperature for 24 hours under hydrogen atmosphere. The suspension was filtered through Celite and the filtrate was dried in vacuo to yield **CD_3-β-D-Man** (17 mg, 82%). Its ^{1}H NMR spectrum was virtually the same as that observed for commercially available methyl-h_3 β -D-mannopyranoside except for the absence of the -OCH_3 signal. ^{1}H NMR (D_2O) δ 4.55 (d, H-1, *J* = 1.0 Hz, 1H), 3.96 (dd, H-2, *J* = 1.0, 3.4 Hz, 1H), 3.91 (dd, H-6a, *J* = 2.4, 12.2 Hz, 1H), 3.71 (dd, H-6b, *J* = 6.8, 12.2 Hz, 1H), 3.61 (dd, H-3, *J* = 3.5, 9.6 Hz, 1H), 3.54 (dd, H-4, *J* = 9.7, 9.7 Hz, 1H), 3.34 (m, H-5, 1H).

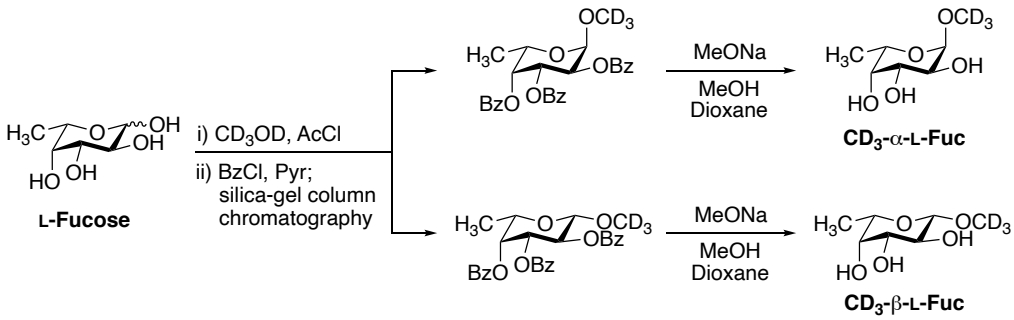
3.5.8 Synthesis of CD_3-α-L-Glc and CD_3-β-L-Glc



CD_3-α-L-Glc **CD_3-β-L-Glc**

CD_3-α-L-Glc and **CD_3-β-L-Glc** were obtained starting from L-glucose in a similar manner to the synthesis of **CD_3-α-D-Glc** and **CD_3-β-D-Glc**.

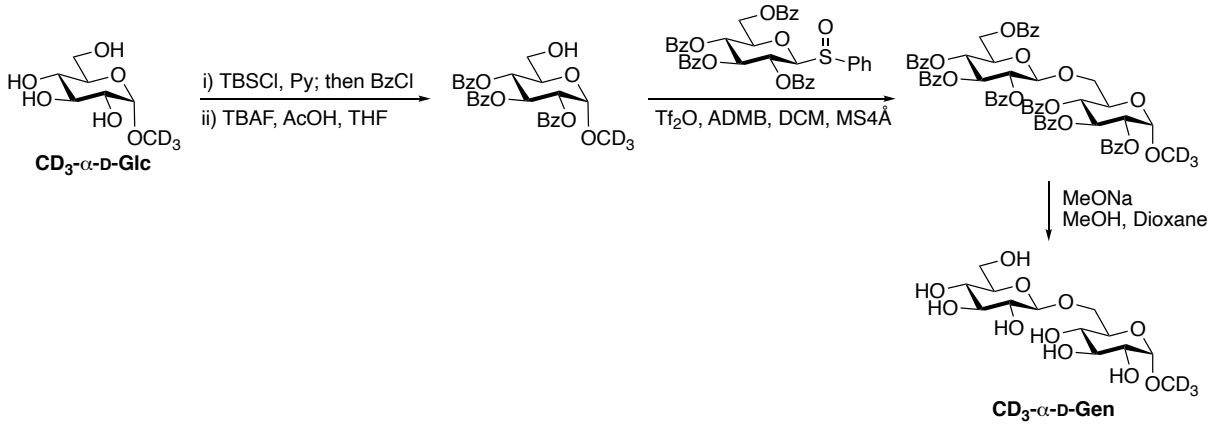
3.5.9 Synthesis of CD_3-α-L-Fuc and CD_3-β-L-Fuc



CD_3-α-L-Fuc and **CD_3-β-L-Fuc** were obtained starting from L-fucose using the general procedures (A)-(C). The ^{1}H NMR spectra of methyl-d_3 2,3,4-tri-*O*-benzoyl-α-D-fucopyranoside, **CD_3-α-L-Ara**, and **CD_3-β-L-Fuc** were virtually the same as those reported for methyl-h_3 2,3,4-tri-*O*-benzoyl-α-D-fucopyranoside[41], methyl-h_3 α-L-fucopyranoside,[42] and methyl-h_3 β-L-fucopyranoside,[43] respectively, except for the absence of the -OCH_3 signal. Methyl-d_3 2,3,4-tri-*O*-benzoyl-α-L-fucopyranoside: ^{1}H NMR ($CDCl_3$) δ 8.13-8.08 (m, ArH, 2H), 8.00-7.95 (m, ArH, 2H), 7.82-7.76 (m, ArH, 2H), 7.61 (m, ArH, 1H), 7.53-7.34 (m, ArH, 6H), 7.28-7.20 (m, ArH, 2H), 5.96 (dd, H-3, *J* = 3.5, 10.7 Hz, 1H), 5.76 (d, H-4, *J* = 2.2 Hz, 1H), 5.64 (dd, H-2, *J* = 3.5, 10.7 Hz, 1H), 5.24 (d, H-1, *J* = 3.5 Hz, 1H), 4.40 (m, H-5, 1H), 1.29 (d, H_3-6, *J* = 6.5 Hz, 3H). Methyl-d_3 2,3,4-tri-*O*-benzoyl-β-L-fucopyranoside: ^{1}H NMR ($CDCl_3$) δ 8.13-8.08 (m, ArH, 2H), 7.99-7.94 (m, ArH, 2H), 7.81-7.75 (m, ArH, 2H), 7.61 (m, ArH, 1H), 7.53-7.35 (m, ArH, 6H), 7.25-7.20 (m, ArH, 2H), 5.78-5.68 (m, H-2, H-4, 2H), 5.56 (dd, H-3, *J* = 3.4, 10.3 Hz, 1H), 4.68 (d, H-1, *J* = 7.9 Hz, 1H), 4.09 (m, H-5, 1H), 1.37 (d, H_3-6, *J* = 6.4 Hz, 3H); ^{13}C NMR ($CDCl_3$) δ 166.2 (C=O), 165.9 (C=O), 165.6 (C=O), 133.6 (Ar), 133.3 (Ar), 133.3 (Ar), 130.1 (Ar), 130.0 (Ar), 129.9 (Ar), 129.4 (Ar), 129.0 (Ar), 129.0 (Ar), 128.7 (Ar), 128.5 (Ar), 128.4 (Ar), 110.9 (C-1), 109.2 (OCD_3), 102.4 (C-2), 72.3 (C-4), 71.2 (C-3), 69.9 (C-5), 16.3 (C-6); $[\alpha]_D$ -206.3 (*c* 1.0, $CHCl_3$); HRMS (ESI) *m/z*

calculated for $C_{28}H_{23}D_3O_8Na$ $[M+Na]^+$ 516.1708, found 516.1693. **CD_3-α-L-Fuc**: ^1H NMR (CD_3OD) δ 4.63 (d, H-1, J = 2.7 Hz, 1H), 3.91 (m, H-5, 1H), 3.75-3.68 (m, H-2, H-3, 2H), 3.65 (s, H-4, 1H), 1.22 (d, H_3-6, J = 6.6 Hz, 3H). **CD_3-β-L-Fuc**: ^1H NMR (D_2O) δ 4.27 (d, H-1, J = 7.8 Hz, 1H), 3.77 (m, H-5, 1H), 3.71 (d, H-4, J = 2.7 Hz, 1H), 3.61 (dd, H-3, J = 3.5, 10.1 Hz, 1H), 3.53-3.41 (m, H-2, 1H), 1.23 (d, H_3-6, J = 6.6 Hz, 3H).

3.5.10 Synthesis of CD_3-α-D-Gen



CD_3-α-D-Glc was added pyridine (5 mL) and TBSCl (224 mg) and stirred overnight at room temperature. To this mixture additional TBSCl (180 mg) was added and stirred at room temperature for 3 hours, which led to the complete consumption of the **CD_3-α-D-Glc** on TLC. To this solution BzCl (1.2 mL) was then added and the mixture was stirred at room temperature for 1.5 hours. The mixture was diluted with Et_2O, washed sequentially with 2 M HCl aq, 2 M NaOH aq, and brine, and then dried over $MgSO_4$. After removal of the solvent, the mixture was added ca. 1 M THF solution of TBAF (3 mL) and AcOH (200 μL) and stirred overnight at room temperature. The mixture was diluted with Et_2O, washed with brine, and dried over $MgSO_4$. After removal of the solvent, the mixture was purified by silica-gel column chromatography (hexane-EtOAc = 2:3 to 1:5), which afforded methyl-d_3 2,3,4-tri-*O*-benzoyl-α-D-glucopyranoside (303 mg, 45% in 3 steps). Its ^1H NMR spectrum was virtually the same as that reported for

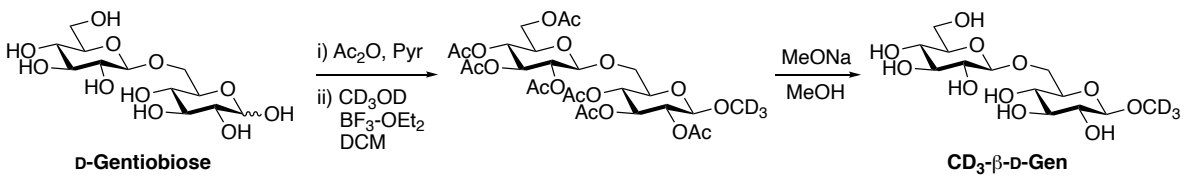
methyl-h_3 2,3,4-tri-*O*-benzoyl-α-D-glucopyranoside except for the absence of the -OCH_3 signal.[44] Methyl-d_3 2,3,4-tri-*O*-benzoyl-α-D-glucopyranoside: ^{1}H NMR ($CDCl_3$) δ 8.00-7.96 (m, ArH, 4H), 7.88 (m, ArH, 2H), 7.53 (m, ArH, 1H), 7.50 (m, ArH, 1H), 7.43-7.35 (m, ArH, 5H), 7.28 (m, ArH, 2H), 6.24 (dd, H-3, *J* = 9.8, 9.8 Hz, 1H), 5.51 (dd, H-4, *J* = 9.9, 9.9 Hz, 1H), 5.30 (dd, H-2, *J* = 3.7, 10.0 Hz, 1H), 5.27 (d, H-1, *J* = 3.7 Hz, 1H), 4.05 (ddd, H-5, *J* = 2.3, 3.9, 10.2 Hz, 1H), 3.84 (dd, H-6a, *J* = 2.3, 13.0 Hz, 1H), 3.75 (dd, H-6b, *J* = 3.8, 12.9 Hz, 1H), 2.78 (br s, OH-6, 1H).

To a 2-neck flask was added methyl-d_3 2,3,4-tri-*O*-benzoyl-α-D-glucopyranoside (48 mg, 94 μmol), phenyl 2,3,4,6-tetra-O-benzoyl-1-sulfinyl-β-D-glucopyranoside[45] (65 mg, 92 μmol), and 4-allyl-1,2-dimethoxybenzene (21 μL). This mixture was dried under vacuum for 1 hour and then added 4 Å molecular sieves (0.3 g). This flask was further dried under vacuum and then purged with N_2. To this flask was added CH_2Cl_2 (2 mL), stirred for 15 mins at room temperature and then cooled to -78 °C. The mixture was then added Tf_2O (16.9 μL) and stirred for 15 mins at this temperature, and then the temperature was slowly raised to -40 °C over 30 mins. The reaction was then quenched with triethylamine (50 μL), diluted with EtOAc, and filtered through cotton. The filtrate was washed sequentially with sat NH_4Cl aq, sat $NaHCO_3$ aq, and brine, and then dried over $MgSO_4$. After removal of the solvent, the mixture was purified by silica-gel column chromatography (hexane-EtOAc = 2:1 to 4:5), which afforded methyl-d_3 2′,3′,4′,6′-tetra-*O*-benzoyl-β-D-glucopyranosyl-(1′→6)-2,3,4-tri-*O*-benzoyl-α-D-glucopyranoside (71 mg, 71%). Its ^{1}H NMR spectrum was virtually the same as that reported for methyl-h_3 2′,3′,4′,6′-tetra-*O*-benzoyl-β-D-glucopyranosyl-(1′→6)-2,3,4-tri-*O*-benzoyl-α-D-glucopyranoside except for the absence of the -OCH_3 signal.[46] Methyl-d_3 2′,3′,4′,6′-tetra-*O*-benzoyl-β-D-glucopyranosyl-(1′→6)-2,3,4-tri-*O*-benzoyl-α-D-glucopyranoside: ^{1}H NMR ($CDCl_3$) δ 8.01-7.77 (m, ArH, 14H), 7.56-7.24 (m, ArH, 21H), 6.06 (dd, H-3, *J* = 9.7, 10.0 Hz, 1H), 5.91 (dd, H-3′, *J* = 9.7, 9.7 Hz, 1H), 5.65 (dd, H-4′, *J* = 9.7, 9.7 Hz,

1H), 5.56 (dd, H-2′, *J* = 7.8, 9.8 Hz, 1H), 5.31 (dd, H-4, *J* = 9.5, 10.3 Hz, 1H), 5.08 (dd, H-2, *J* = 3.6, 10.2 Hz, 1H), 4.97 (d, H-1′, *J* = 7.9 Hz, 1H), 4.93 (d, H-1, *J* = 3.6 Hz, 1H), 4.60 (dd, H-6′a, *J* = 3.2, 12.2 Hz, 1H), 4.44 (dd, H-6′b, *J* = 5.1, 12.1 Hz, 1H), 4.22 (ddd, H-5, *J* = 1.9, 7.6, 10.0 Hz, 1H), 4.14 (ddd, H-5′, *J* = 3.3, 5.1, 9.8 Hz, 1H), 4.10 (dd, H-6a, *J* = 2.0, 11.5 Hz, 1H), 3.78 (dd, H-6b, *J* = 7.8, 11.4 Hz, 1H).

Using the general procedure (C), methyl-d_3 2′,3′,4′,6′-tetra-*O*-benzoyl-β-D-glucopyranosyl-(1′→6)-2,3,4-tri-*O*-benzoyl-α-D-glucopyranoside (51 mg, 47 μmol) was converted to **CD₃-α-D-Gen** (17 mg, quant). Its ^{1}H NMR spectrum was virtually the same as that reported for methyl-h_3 β-D-glucopyranosyl-(1′→6)-α-D-glucopyranoside except for the absence of the -OCH_3 signal.[47] **CD₃-α-D-Gen**: ^{1}H NMR (D_2O) δ 4.78 (d, H-1, *J* = 3.8 Hz, 1H), 4.47 (d, H-1′, *J* = 7.9 Hz, 1H), 4.14 (dd, H-6a, *J* = 2.1, 11.5 Hz, 1H), 3.91-3.85 (m, H-6b, H-6′a, 2H), 3.77 (ddd, H-5, *J* = 2.1, 5.0, 10.1 Hz, 1H), 3.70 (dd, H-6□b, *J* = 5.9, 12.3 Hz, 1H), 3.64 (dd, H-3, *J* = 9.5, 9.5 Hz, 1H), 3.54 (dd, H-2, *J* = 3.8, 9.8 Hz, 1H), 3.49 (dd, H-3′, *J* = 9.1, 10.1 Hz, 1H), 3.46 (dd, H-4, *J* = 9.1, 9.1 Hz, 1H), 3.43 (ddd, H-5′, *J* = 2.3, 5.9, 9.9 Hz, 1H), 3.36 (dd, H-4′, *J* = 9.0, 9.8 Hz, 1H), 3.28 (dd, H-2′, *J* = 7.9, 9.4 Hz, 1H).

3.5.11 Synthesis of CD₃-β-D-Gen



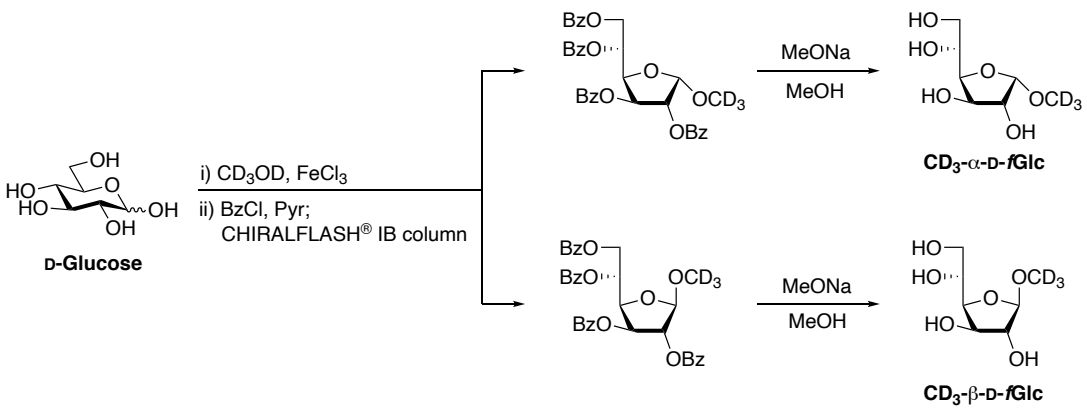
D-gentiobiose (118 mg, 345 μmol) was added pyridine (4 mL) and acetic anhydride (2 mL) at 80 °C, and stirred overnight at 50 °C. The mixture was then dried in vacuo, which provided pure D-gentiobiose octaacetate (234 mg, quant). Part of this product (224 mg,

330 μmol) was dissolved in CH_2Cl_2 (3 mL), added CD_3OD (54 μL) and BF_3-OEt_2 (103 μL) at 0 °C, and stirred overnight at room temperature. The mixture was diluted with EtOAc, washed sequentially with sat $NaHCO_3$ aq and brine, and then dried over $MgSO_4$. After removal of the solvent, the mixture was purified by silica-gel column chromatography (hexane-EtOAc = 1:1 to 1:4), which provided methyl-d_3 2′,3′,4′,6′-tetra-*O*-acetyl-β-D-glucopyranosyl-(1′→6)-2,3,4-tri-*O*-acetyl-β-D-glucopyranoside (87 mg, 40%). Its ^{1}H NMR spectrum was virtually the same as that reported for methyl-h_3 2′,3′,4′,6′-tetra-*O*-acetyl-β-D-glucopyranosyl-(1′→6)-2,3,4-tri-*O*-acetyl-β-D-glucopyranoside except for the absence of the -OCH_3 signal.[48] Methyl-d_3 2′,3′,4′,6′-tetra-*O*-acetyl-β-D-glucopyranosyl-(1′→6)-2,3,4-tri-*O*-acetyl-β-D-glucopyranoside: ^{1}H NMR ($CDCl_3$) δ 5.19 (dd, H-3, *J* = 9.5, 9.5 Hz, 1H), 5.18 (dd, H-3′, *J* = 9.5, 9.5 Hz, 1H), 5.07 (dd, H-4′, *J* = 9.7, 9.7 Hz, 1H), 5.00 (dd, H-2′, *J* = 8.0, 9.6 Hz, 1H), 4.93 (dd, H-2, *J* = 8.0, 9.7 Hz, 1H), 4.89 (dd, H-4, *J* = 9.6, 9.6 Hz, 1H), 4.60 (d, H-1′, *J* = 7.9 Hz, 1H), 4.39 (d, H-1, *J* = 7.9 Hz, 1H), 4.27 (dd, H-6′a, *J* = 4.8, 12.3 Hz, 1H), 4.13 (dd, H-6′b, *J* = 2.4, 12.3 Hz, 1H), 3.88 (dd, H-6a, *J* = 2.0, 11.0 Hz, 1H), 3.72–3.66 (m, H-5, H-5′, 2H), 3.62 (dd, H-6b, *J* = 7.5, 11.0 Hz, 1H), 2.09 (s, Ac, 3H), 2.05 (s, Ac, 3H), 2.04 (s, Ac, 3H), 2.03 (s, Ac, 3H), 2.02 (s, Ac, 3H), 2.00 (s, Ac, 3H), 1.99 (s, Ac, 3H).

In a similar manner to the general procedure (C) without the use of 1,4-dioxane, methyl-d_3 2′,3′,4′,6′-tetra-*O*-acetyl-β-D-glucopyranosyl-(1′→6)-2,3,4-tri-*O*-acetyl-β-D-glucopyranoside (87 mg, 133 μmol) was converted to **CD_3-β-D-Gen** (25 mg, 52%). Its ^{1}H NMR spectrum was virtually the same as that reported for methyl-h_3 β-D-glucopyranosyl-(1′→6)-β-D-glucopyranoside except for the absence of the -OCH_3 signal.[47] **CD_3-β-D-Gen**: ^{1}H NMR (D_2O) δ 4.49 (d, H-1′, *J* = 7.9 Hz, 1H), 4.36 (d, H-1, *J* = 8.0 Hz, 1H), 4.19 (dd, H-6a, *J* = 2.1, 11.7 Hz, 1H), 3.90 (dd, H-6′a, *J* = 2.3, 12.3 Hz, 1H), 3.84 (dd, H-6b, *J* = 5.8, 11.7 Hz, 1H), 3.70 (dd, H-6′b, *J* = 5.8, 12.3 Hz, 1H), 3.59

(ddd, H-5, J = 2.0, 5.8, 9.7 Hz, 1H), 3.50-3.40 (m, H-3, H-4, H-3′, H-5′, 4H), 3.37 (dd, H-4′, J = 8.9, 9.8 Hz, 1H), 3.29 (dd, H-2′, J = 8.0, 9.4 Hz, 1H), 3.24 (dd, H-2, J = 8.0, 9.2 Hz, 1H).

3.5.12 Synthesis of CD_3-α-D-*f*Glc and CD_3-β-D-*f*Glc



CD_3-α-D-*f*Glc and **CD_3-β-D-*f*Glc** were synthesized in a similar manner to a reported procedure.[30] A solution of $FeCl_3$ (389 mg, 2.40 mmol) in CD_3OD (10 mL) was added D-glucose (324 mg, 1.80 mmol) and stirred at 40 °C for 2 days. This mixture was added Celite and then sat $NaHCO_3$ aq (10 mL) and then filtered through a pad of Celite, which was washed with MeOH. After removal of the solvent, the residue was dissolved in THF, and the solution was filtered and evaporated. The crude mixture was purified by silica-gel column chromatography ($CHCl_3$-MeOH = 7:1 to 3:1), which provided a mixture of methyl-d_3 α- and β-D-glucofuranosides (225 mg, 63%, α/β = 3:2). The mixture (200 mg, 1.01 mmol) was used for the next reaction using the general procedures (B) to produce the tetrabenzoyl synthetic intermediates (375 mg, 62%, α/β = 2:3).

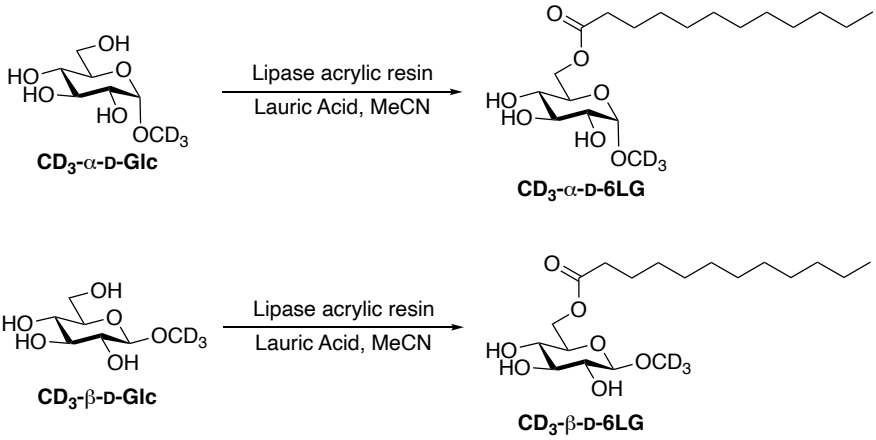
The enantioseperation of tetrabenzoyl synthetic intermediates was carried out on a Daicel CHIRALPAK® IB column (10 mm ϕ × 250 mm) using hexane-EtOAc = 88:12, which led to the first-eluted methyl-d_3 2,3,4,6-tetra-*O*-benzoyl-α-D-glucofuranoside and

the second-eluted methyl-d_3 2,3,4,6-tetra-*O*-benzoyl-β-D-glucofuranoside. Methyl-d_3 2,3,4,6-tetra-*O*-benzoyl-α-D-glucofuranoside: ^{1}H NMR ($CDCl_3$) δ 8.08 (d, ArH, J = 7.3 Hz, 2H), 8.01 (d, ArH, J = 7.3 Hz, 2H), 7.84 (d, ArH, J = 7.3 Hz, 2H), 7.77 (d, ArH, J = 7.3 Hz, 2H), 7.59-7.52 (m, ArH, 2H), 7.50-7.39 (m, ArH, 6H), 7.33-7.27 (m, ArH, 4H), 6.16 (dd, H-3, J = 4.9, 5.7 Hz, 1H), 5.79 (m, H-5, 1H), 5.41 (d, H-1, J = 4.6 Hz, 1H), 5.37 (dd, H-2, J = 4.5, 4.5 Hz, 1H), 4.93-4.86 (m, H-4, H-6a, 2H), 4.71 (dd, H-6b, J = 5.1, 12.5 Hz, 1H); ^{13}C NMR ($CDCl_3$) δ 166.1 (C=O), 165.8 (C=O), 165.3 (C=O), 165.1 (C=O), 133.4 (Ar), 133.3 (Ar), 133.1 (Ar), 133.0 (Ar), 130.0 (Ar), 129.8 (Ar), 129.7 (Ar), 129.5 (Ar), 129.3 (Ar), 129.1 (Ar), 128.6 (Ar), 128.4 (Ar), 128.4 (Ar), 128.2 (Ar), 128.2 (Ar), 102.1 (OCD_3), 100.7 (C-1), 78.7 (C-2), 75.0 (C-3), 73.4 (C-4), 69.6 (C-5), 63.7 (C-6); $[\alpha]_D$ +45.9 (*c* 1.0, $CHCl_3$); HRMS (ESI) *m/z* calcd for $C_{35}H_{27}D_3O_{10}Na$ $[M+Na]^+$ 636.1920, found 636.1907. Methyl-d_3 2,3,4,6-tetra-*O*-benzoyl-β-D-glucofuranoside: ^{1}H NMR ($CDCl_3$) δ 8.07 (d, ArH, J = 7.3 Hz, 2H), 8.02 (d, ArH, J = 6.8 Hz, 2H), 7.86 (d, ArH, J = 7.3 Hz, 2H), 7.75 (d, ArH, J = 7.3 Hz, 2H), 7.63-7.58 (m, ArH, 1H), 7.57-7.52 (m, ArH, 1H), 7.51-7.39 (m, ArH, 6H), 7.34-7.26 (m, ArH, 4H), 5.93 (d, H-3, J = 5.4 Hz, 1H), 5.81 (m, H-5, 1H), 5.48 (s, H-2, 1H), 5.16 (s, H-1, 1H), 5.04 (dd, H-4, J = 5.6, 9.0 Hz, 1H), 4.98 (dd, H-6a, J = 2.0, 12.2 Hz, 1H), 4.73 (dd, H-6b, J = 4.6, 12.5 Hz, 1H); ^{13}C NMR ($CDCl_3$) δ 166.1 (C=O), 165.1 (C=O), 165.0 (C=O), 165.0 (C=O), 133.6 (Ar), 133.3 (Ar), 133.0 (Ar), 133.0 (Ar), 129.9 (Ar), 129.9 (Ar), 129.8 (Ar), 129.7 (Ar), 129.5 (Ar), 129.3 (Ar), 129.0 (Ar), 128.7 (Ar), 128.5 (Ar), 128.4 (Ar), 128.2 (Ar), 128.1 (Ar), 107.5 (C-1), 100.3 (OCD_3), 81.0 (C-2), 78.4 (C-4), 74.6 (C-3), 70.2 (C-5), 64.0 (C-6); $[\alpha]_D$ -39.5 (*c* 1.0, $CHCl_3$); HRMS (ESI) *m/z* calcd for $C_{35}H_{27}D_3O_{10}Na$ $[M+Na]^+$ 636.1920, found 636.1902.

In a similar manner to the general procedure (C) without the use of 1,4-dioxane, methyl-d_3 2,3,4,6-tetra-*O*-benzoyl-α-D-glucofuranoside (52 mg, 87 μmol) and methyl-d_3 2,3,4,6-tetra-*O*-benzoyl-β-D-glucofuranoside (57 mg, 95 μmol) were converted to **CD_3-**

α-D-*f*Glc (10 mg, 58%) and **CD_3-β-D-*f*Glc** (12 mg, 63%), respectively. Their ^{1}H NMR spectra were virtually the same as those reported for methyl-h_3 α- and β-D-glucofuranoside except for the absence of the -OCH_3 signal.[49] **CD_3-α-D-*f*Glc** ^{1}H NMR (D_2O) δ 5.06 (d, H-1, *J* = 4.4 Hz, 1H), 4.27 (dd, H-3, *J* = 3.4, 4.7 Hz, 1H), 4.14 (dd, H-2, *J* = 3.7, 3.7 Hz, 1H), 4.08 (dd, H-4, *J* = 4.6, 8.1 Hz, 1H), 3.86 (m, H-5, 1H), 3.78 (dd, H-6a, *J*= 2.7, 12.0 Hz, 1H), 3.64 (dd, H-6b, *J* = 6.4, 12.2 Hz, 1H). **CD_3-β-D-*f*Glc**: ^{1}H NMR (D_2O) δ 4.86 (s, H-1, 1H), 4.22 (d, H-3, *J* = 4.4 Hz, 1H), 4.16 (dd, H-4, *J* = 4.6, 9.0 Hz, 1H), 4.12 (s, H-2, 1H), 3.94 (m, H-5, 1H), 3.85 (dd, H-6a, *J* = 2.8, 12.0 Hz, 1H), 3.69 (dd, *J* = 6.1, 12.0 Hz, H-6b, 1H).

3.5.13 Enzymatic preparation of CD_3-α-D-6LG and CD_3-β-D-6LG



CD_3-α-D-6LG was synthesized in a similar manner to a reported procedure.[33] A solution of **CD_3-α-D-Glc** (97 mg, 0.50 mmol) in MeCN (1.5 mL), was added lauric acid (100 mg), thermally activated pulverized 4 Å molecular sieves (80 mg) and lipase acrylic resin (≥5,000 U/g, recombinant, expressed in *Aspergillus niger*) purchased from Sigma-Aldrich (0.49g, 50% w/w of sugar). The reaction mixture was incubated in a shaking incubator S1-300C (from AS ONE) at 60 °C with rotational speed of 300 rpm for 3 days.

The reaction mixture was diluted with EtOAc and then filtered, and the filtrate was evaporated. The residue was purified by silica-gel column chromatography ($CHCl_3$-MeOH = 15:1 to 9:1), which provided **CD_3-α-D-6LG**, (53 mg, 28%). Its ^{1}H NMR spectrum was virtually the same as that reported for methyl-h_3 6-lauroyl-α-D-glucopyranoside except for the absence of the -OCH_3 signal. [32] **CD_3-α-D-6LG**: ^{1}H NMR ($CDCl_3$) δ 4.79 (d, H-1, J = 3.9 Hz, 1H), 4.57 (dd, H-6a, J = 4.2, 12.5, Hz, 1H), 4.22 (dd, H-6b, J = 2.0, 12.2 Hz, 1H), 3.76-3.69 (m, H-3, H-5, 2H), 3.53 (dd, H-2, J = 3.9, 9.3 Hz, 1H), 3.34 (dd, H-4, J = 9.3, 9.3, Hz, 1H), 2.38 (t, CH_2, J = 7.6 Hz, 2H), 1.64 (m, CH_2, 2H), 1.22-1.35 (m, CH_2 × 8, 16H), 0.88 (t, CH_3, J = 6.8 Hz, 3H).

CD_3-β-D-Glc (89 mg, 0.45 mmol) was converted to **CD_3-β-D-6LG** (60 mg, 32%) in a similar manner. Its ^{1}H NMR spectrum was similar to that reported for methyl-h_3 6-lauroyl-β-D-glucopyranoside, but the signal assignment and some of chemical shifts were slightly different from ours.[32] **CD_3-β-D-6LG**: ^{1}H NMR ($CDCl_3$) δ 4.59 (dd, H-6a, J = 3.9, 12.2 Hz, 1H), 4.26 (dd, H-6b, J = 2.0, 12.2 Hz, 1H), 4.21 (d, H-1, J = 7.8 Hz, 1H), 3.59 (dd, H-3, J = 9.1, 9.1, 1H), 3.45 (m, H-5, 1H), 3.41-3.34 (m, H-2, H-4, 2H), 2.38 (t, CH_2, J = 7.6 Hz, 2H), 1.63 (m, CH_2, 2H), 1.22-1.34 (m, CH_2 × 8, 16H), 0.88 (t, CH_3, J = 6.8 Hz, 3H).

3.6 Supporting Informations

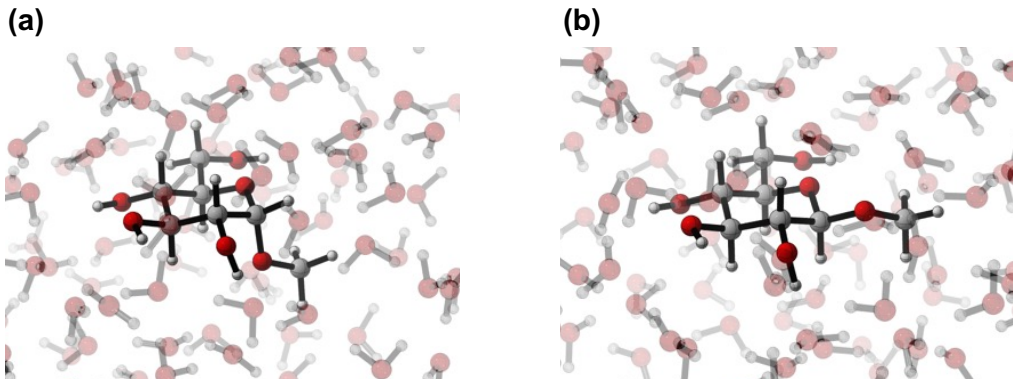


Figure S1 Representative MD snapshots for (a) **CD_3-α-D-Glc** and (b) **CD_3-β-D-Glc** with explicit water molecules.

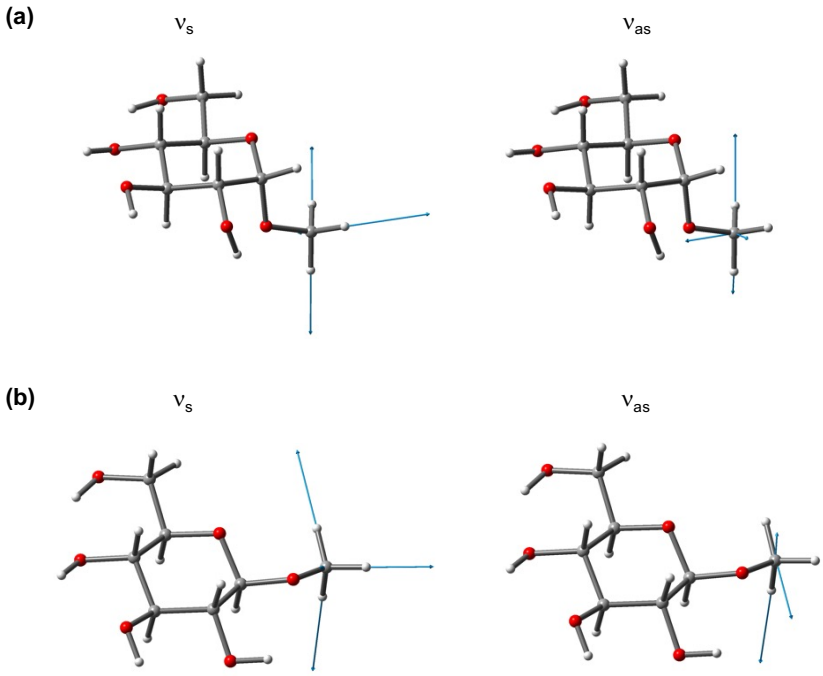


Figure S2 Vibrational modes of ν_s (left) and ν_{as} (right) CD_3 group of (a) **CD_3-α-D-Glc** and (b) **CD_3-β-D-Glc** calculated at DFT/B3LYP/6-31G(d)/PCM(water). Displacement vectors (arbitrary length) of each atom are shown as arrows.

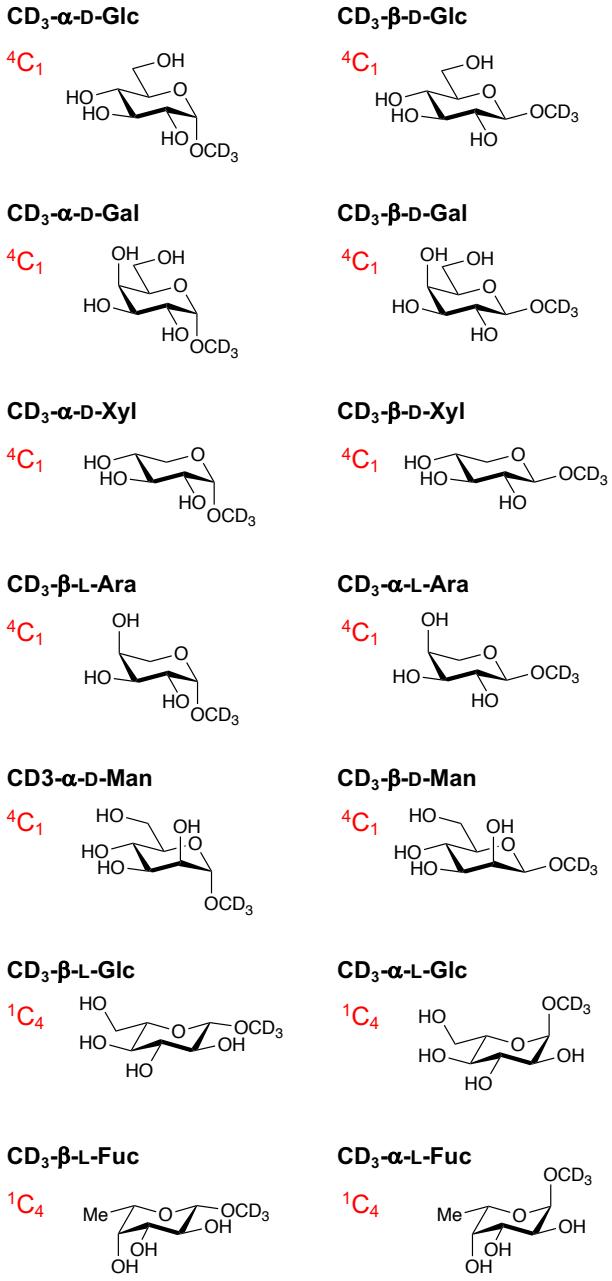


Figure S3 Preferred pyranose conformations of the studied methyl-d_3 glycopyranosides in polar solvents.

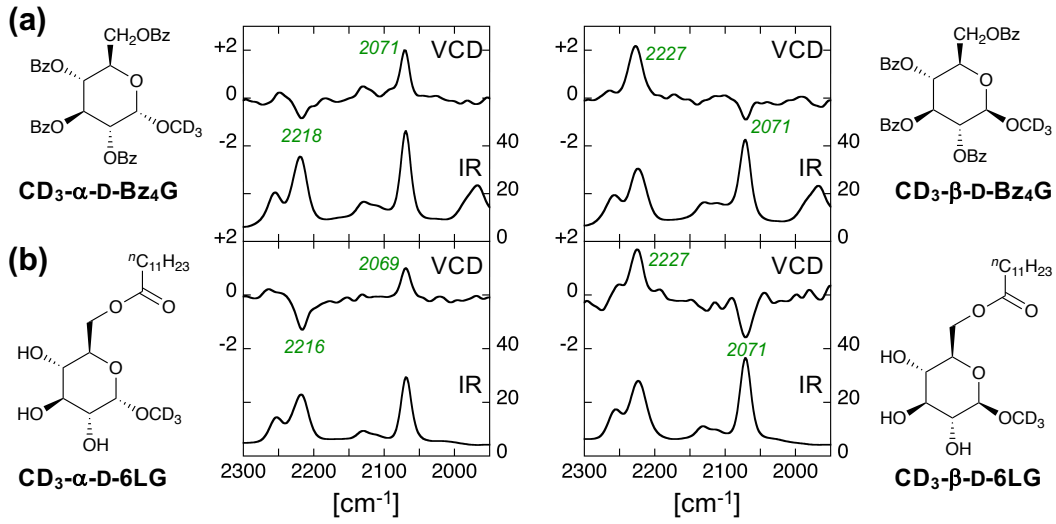
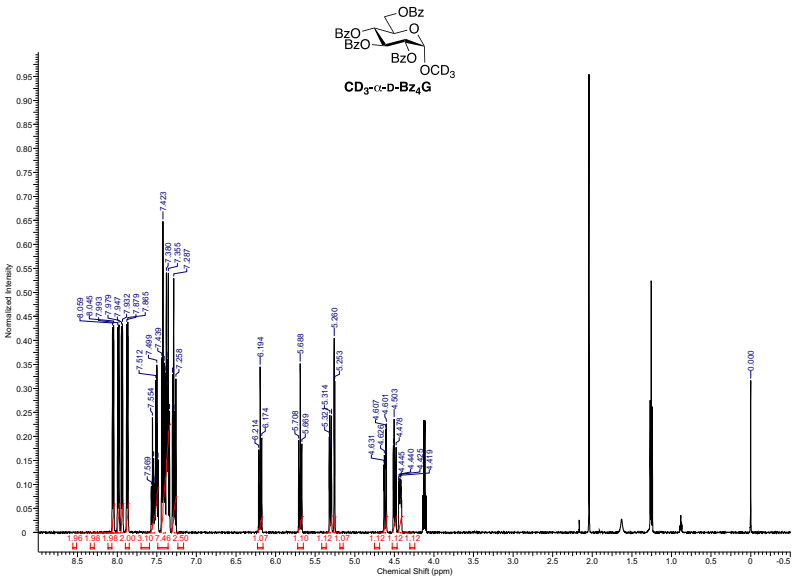
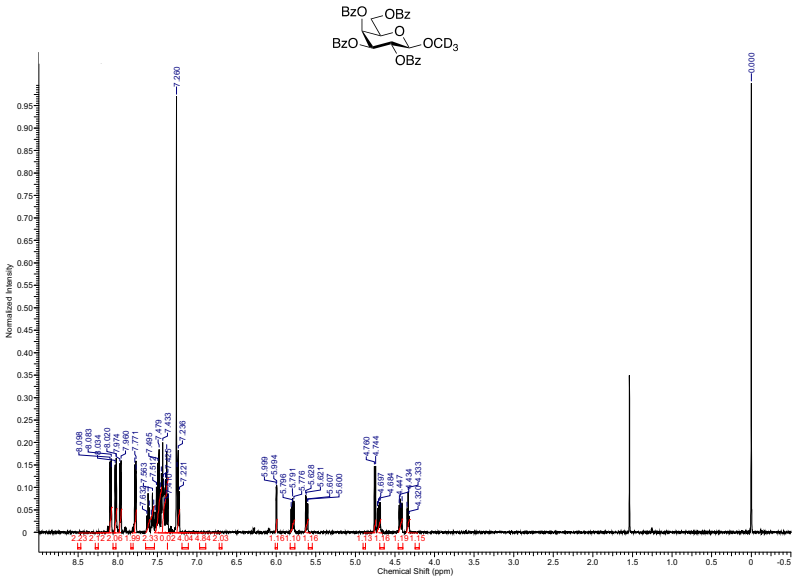
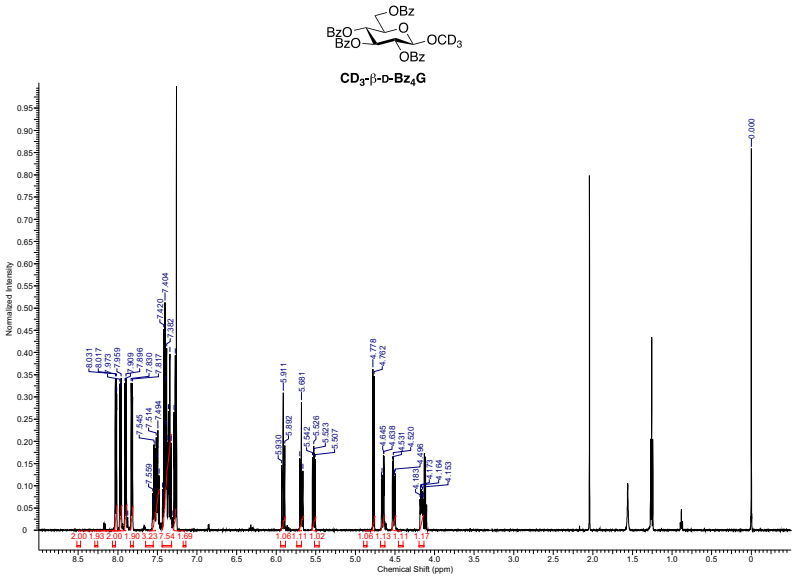


Figure S4 VCD and IR spectra of (a) **CD_3-α-** and **β-D-Bz_4G** and (b) **CD_3-α-** and **β-D-6LG** measured in $CHCl_3$ (*c* 0.5 or 1.0 M, *l* 50 μm). Wavenumbers of VCD signal extrema for ν_s CD_3 (lower frequency) and ν_{as} CD_3 (higher frequency) are labelled in italic (green).

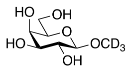
NMR Spectra



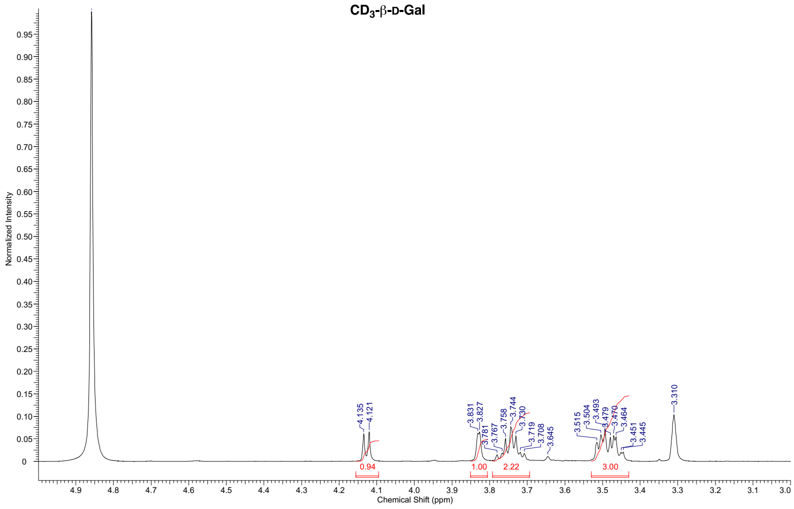

BzO
OBz
BzO
O
OCD3
OBz
Normalized Intensity
Chemical Shift (ppm)

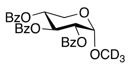

OBz
BzO
BzO
O
OCD3
OBz
Normalized Intensity
Chemical Shift (ppm)

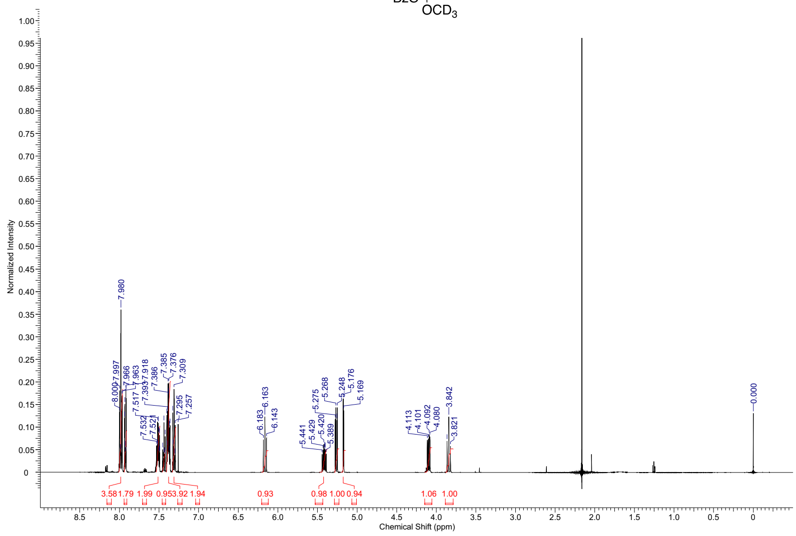
CD_3-β-D-Bz_4G

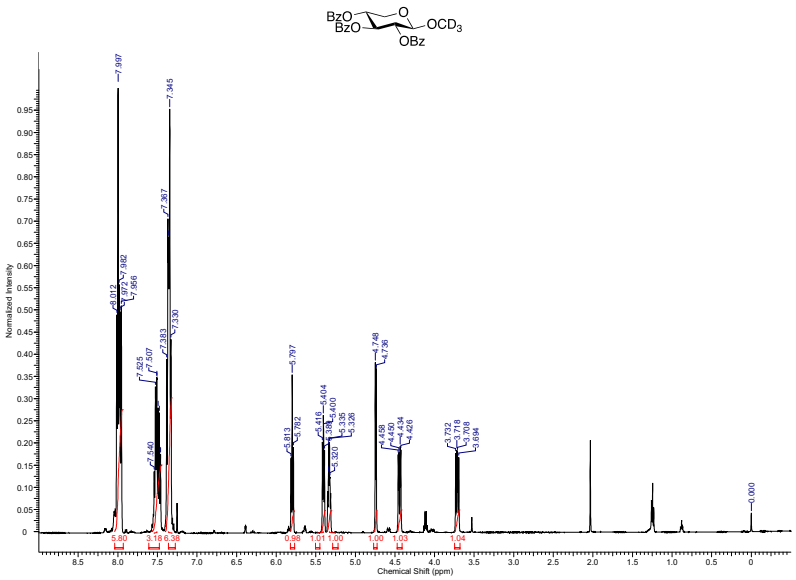

HO
OH
O
HO
OCD3
OH

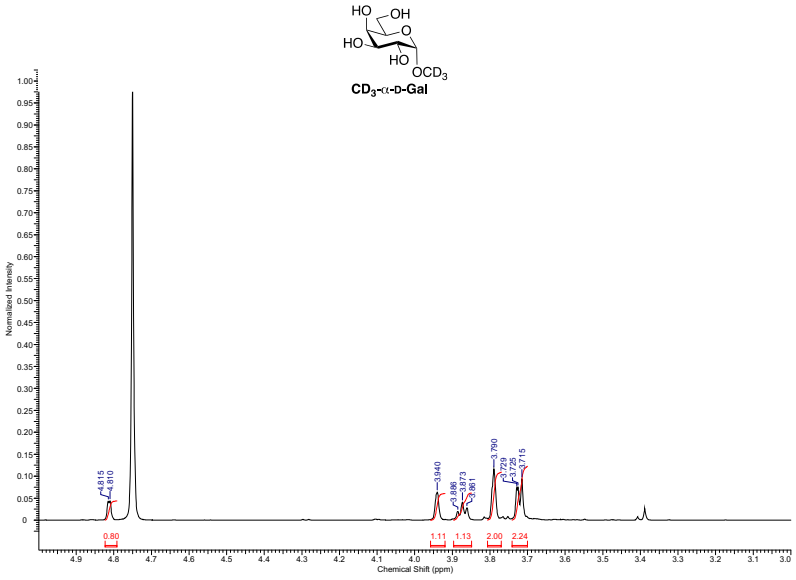
CD_3-β-D-Gal


Normalized Intensity
Chemical Shift (ppm)

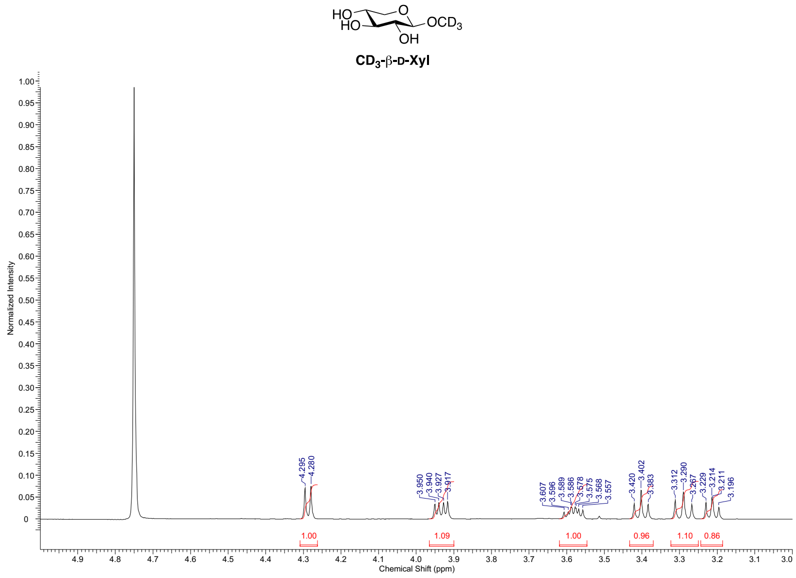

BzO
O
BzO
BzO
OCD3

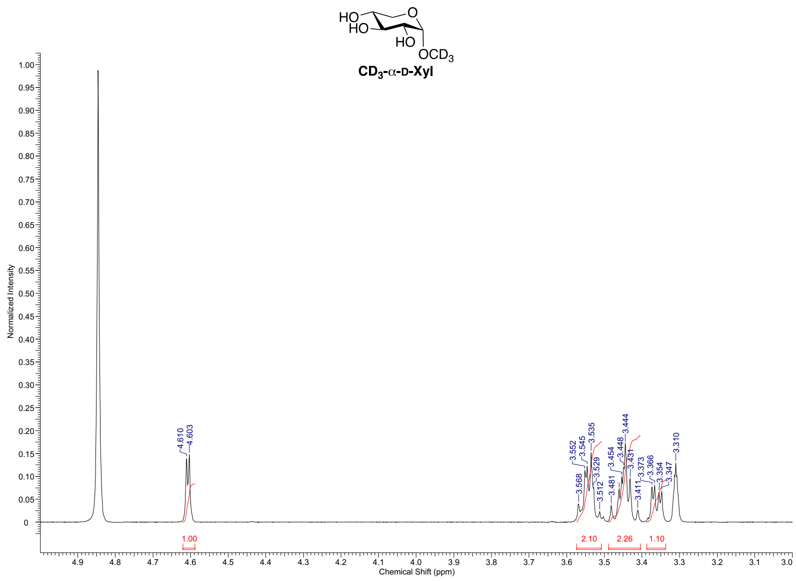

Normalized Intensity
Chemical Shift (ppm)

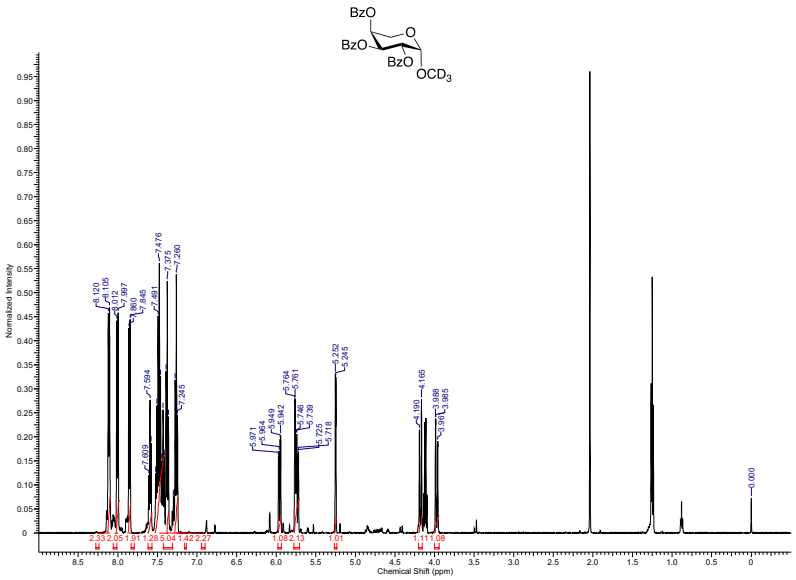

BzO
BzO
O
OCD3
OBz
Normalized Intensity
Chemical Shift (ppm)

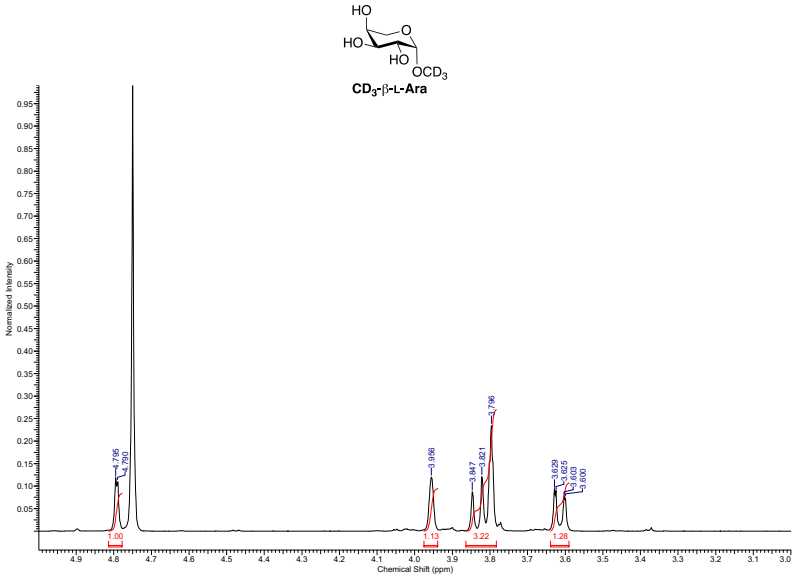

HO
OH
O
HO
HO
OCD3
Normalized Intensity
Chemical Shift (ppm)

CD_3-α-D-Gal

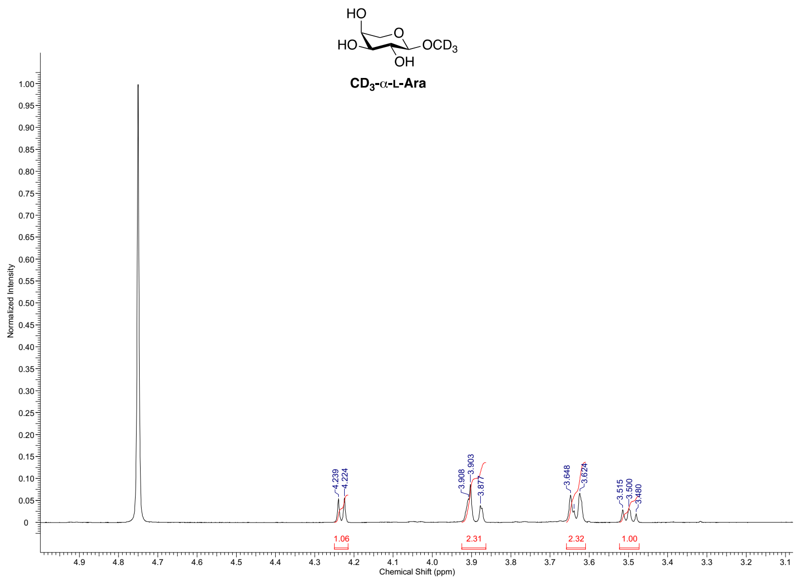

HO
HO
O
OCD3
OH
CD3-β-D-Xyl
Normalized Intensity
Chemical Shift (ppm)
4.295
4.280
3.950
3.940
3.927
3.917
3.607
3.596
3.589
3.586
3.578
3.575
3.568
3.557
3.420
3.402
3.383
3.312
3.290
3.267
3.229
3.214
3.211
3.196
1.00
1.09
1.00
0.96
1.10
0.86


HO
HO
O
HO
OCD3
CD3-α-D-Xyl
Normalized Intensity
Chemical Shift (ppm)
4.610
4.603
3.568
3.552
3.545
3.535
3.529
3.512
3.481
3.454
3.448
3.444
3.431
3.411
3.373
3.366
3.354
3.347
3.310
1.00
2.10
2.26
1.10

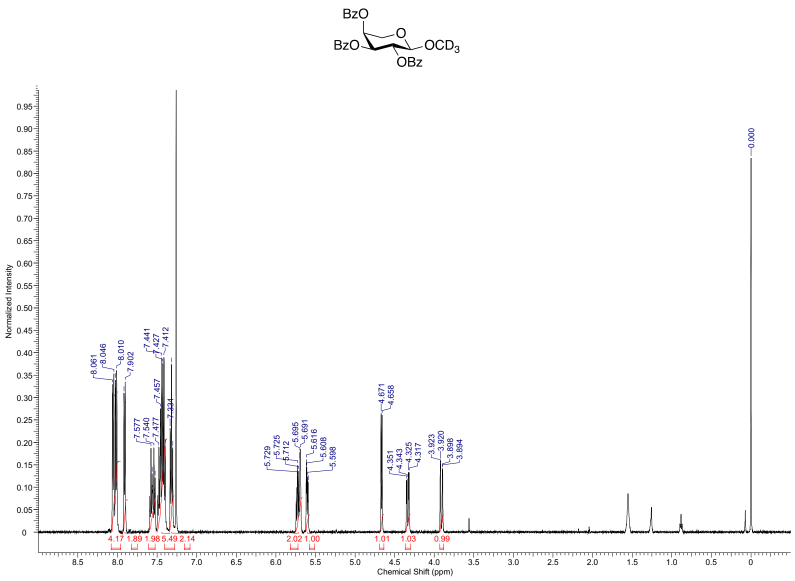

BzO
O
BzO
BzO
OCD3
Normalized Intensity
Chemical Shift (ppm)

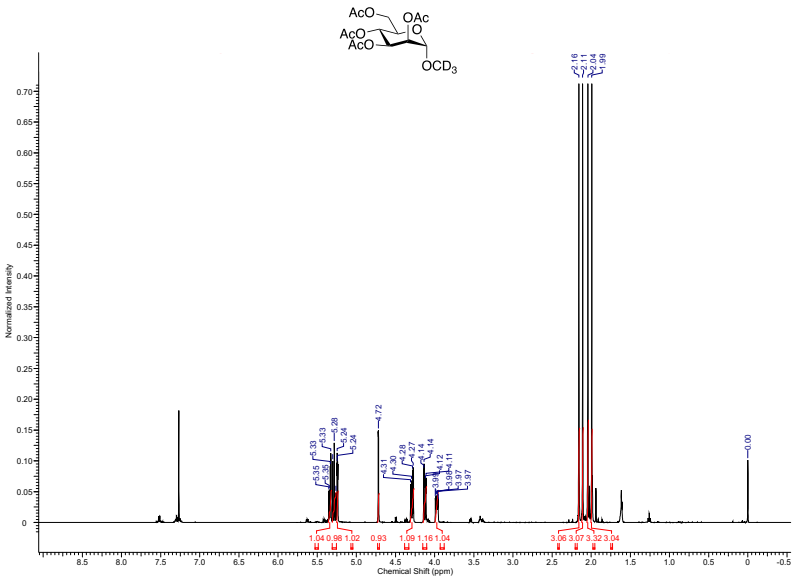

HO
O
HO
HO
OCD3
Normalized Intensity
Chemical Shift (ppm)

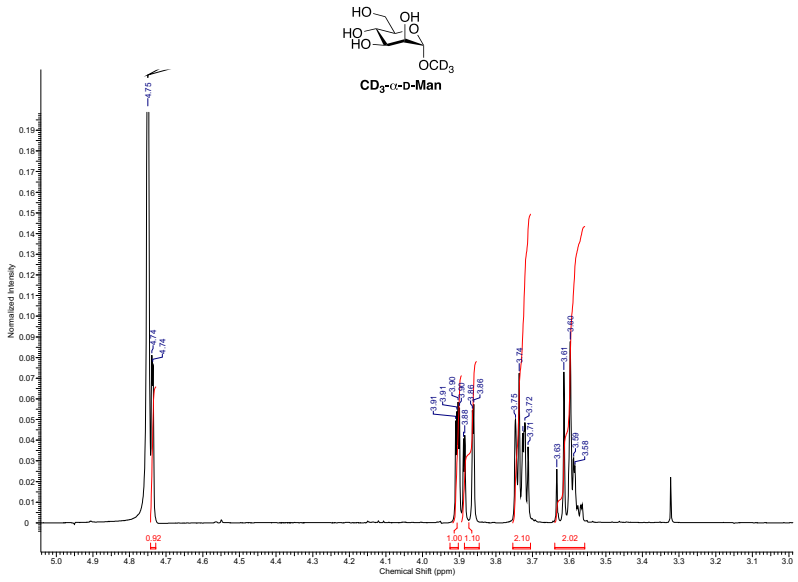
CD_3-β-L-Ara

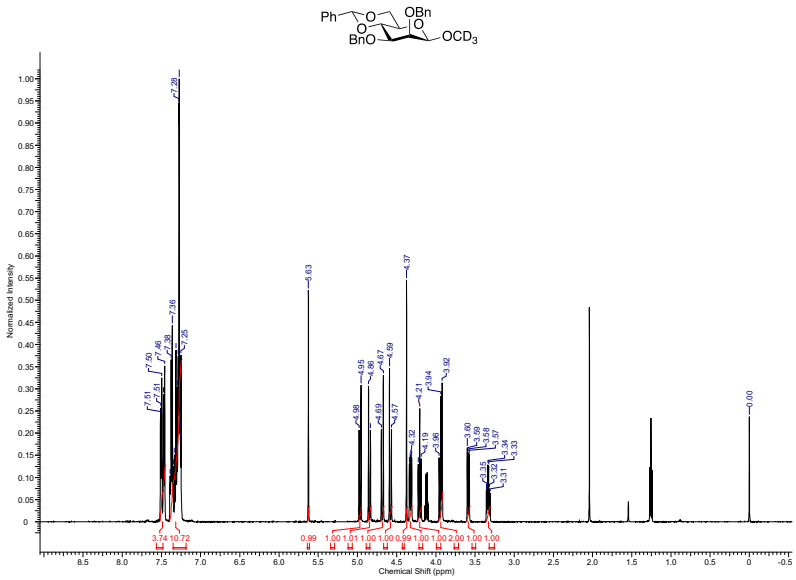


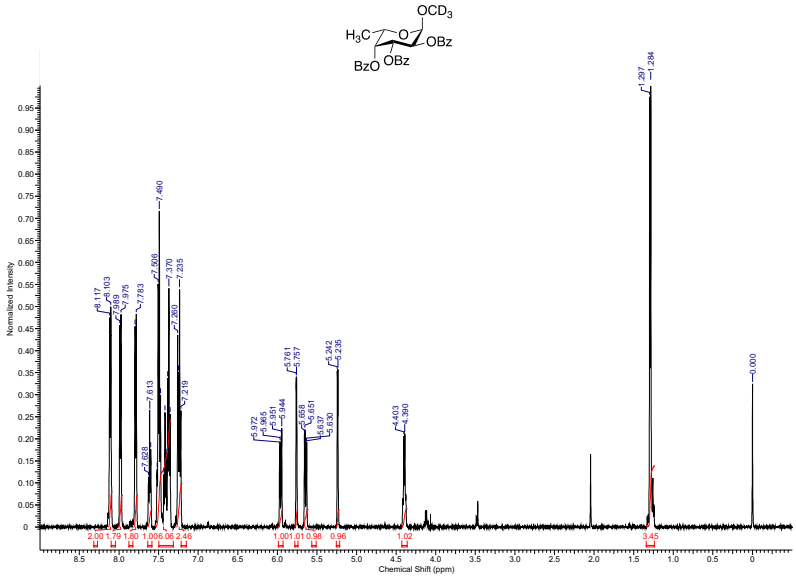
CD_3-α-L-Ara

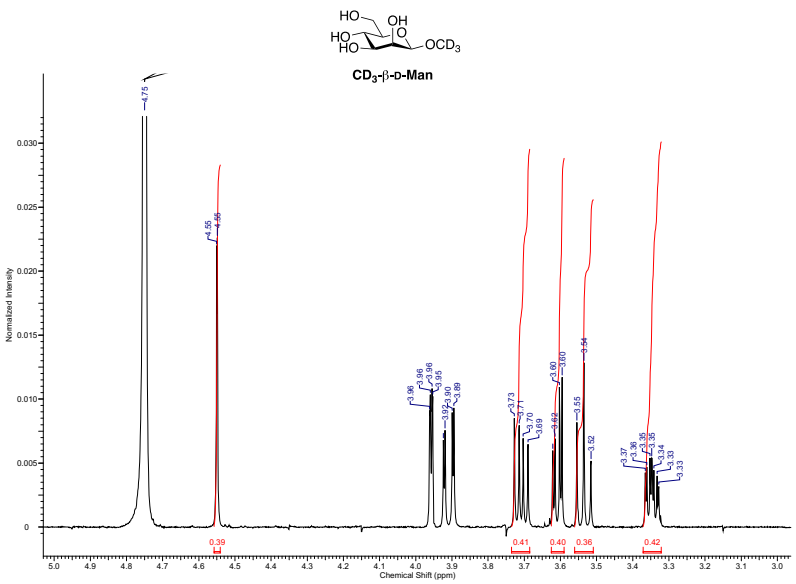


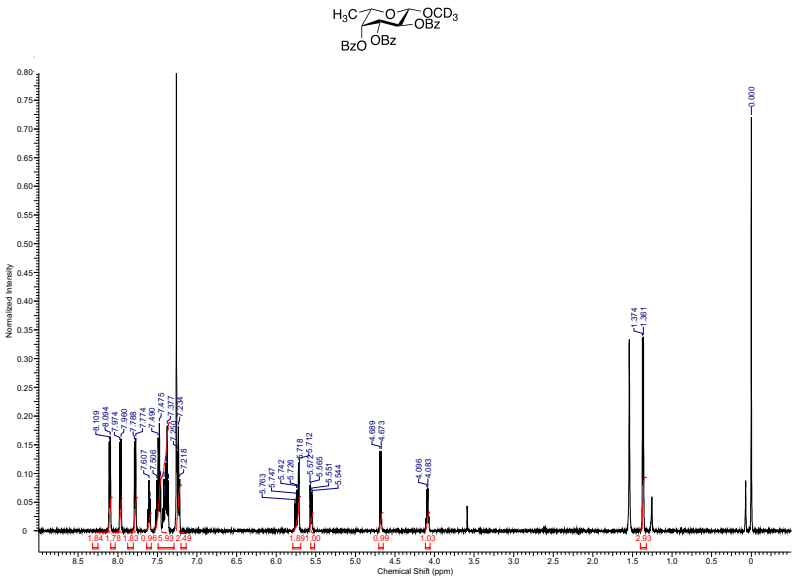

AcO
OAc
AcO
AcO
O
OCD3
Normalized Intensity
Chemical Shift (ppm)

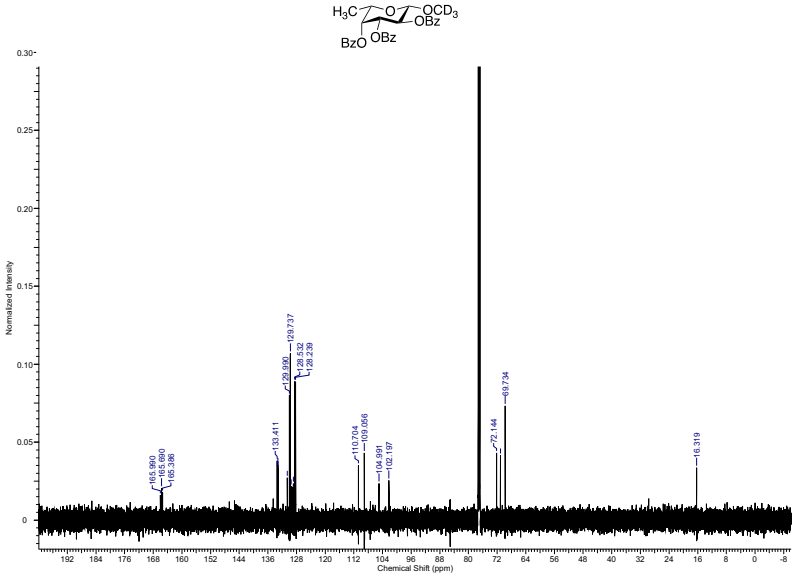

HO
OH
HO
HO
O
OCD3
CD3-α-D-Man
Normalized Intensity
Chemical Shift (ppm)

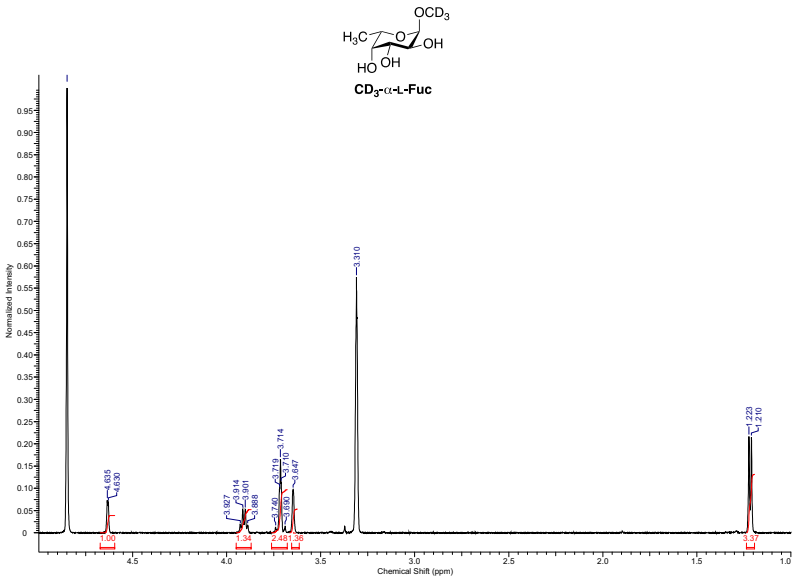

Ph
O
OBn
O
BnO
OCD3
Normalized Intensity
Chemical Shift (ppm)

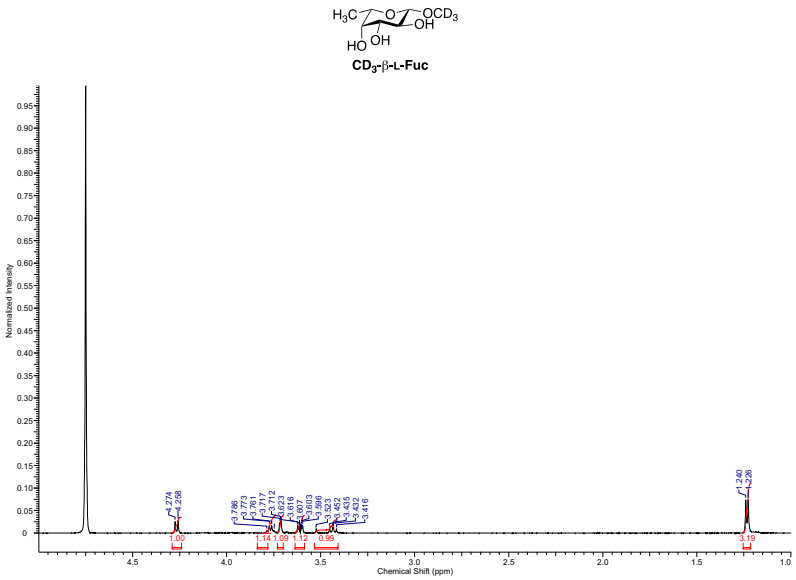

OCD3
H3C
O
OBz
BzO
OBz
Normalized Intensity
Chemical Shift (ppm)

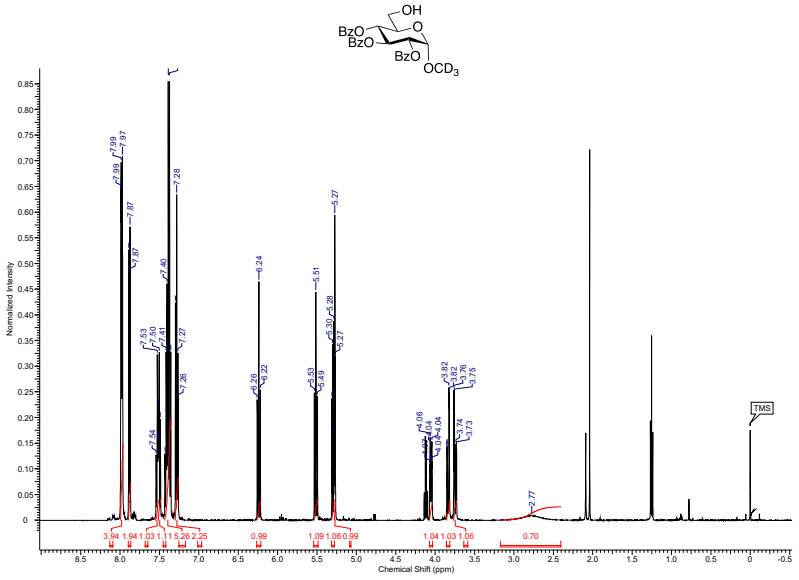

HO
OH
HO
O
HO
OCD3
CD3-β-D-Man
Normalized Intensity
Chemical Shift (ppm)

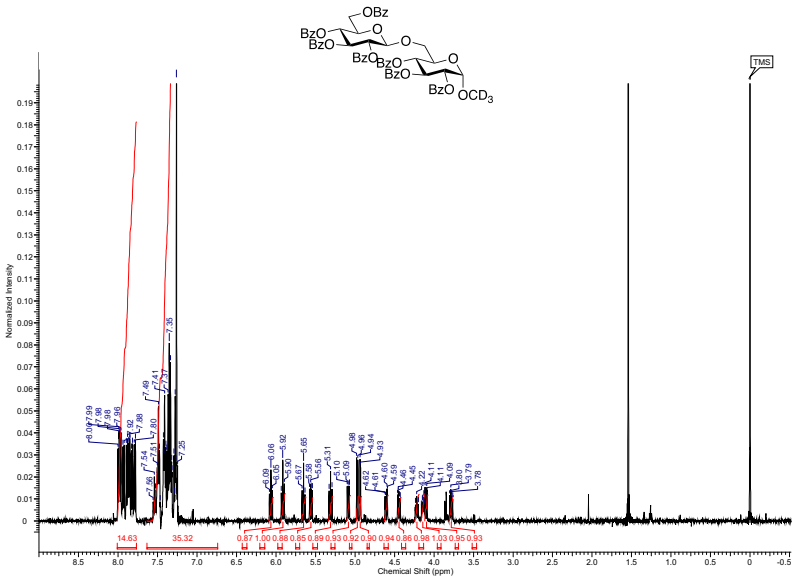

H3C
O
OCD3
OBz
BzO
OBz
Normalized Intensity
Chemical Shift (ppm)

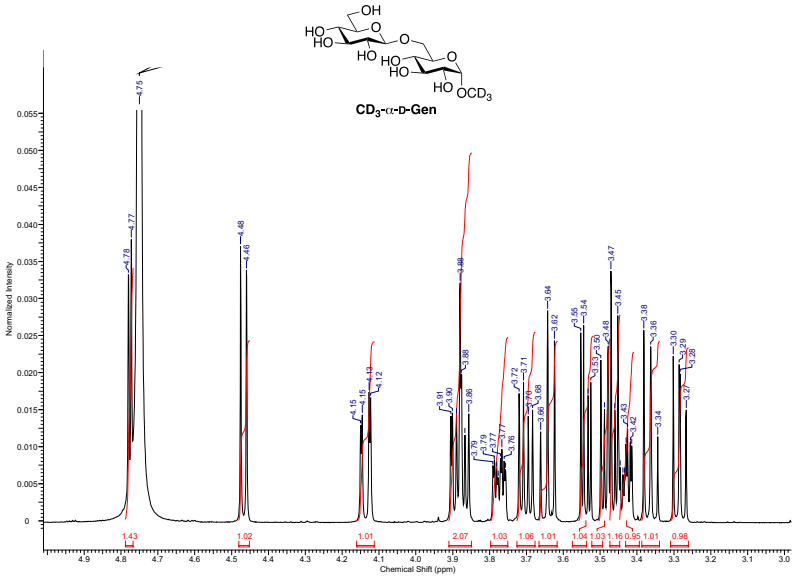

H3C
O
OCD3
OBz
BzO
OBz
Normalized Intensity
Chemical Shift (ppm)
165.990
165.690
165.386
133.411
129.880
129.737
128.532
128.239
110.704
109.956
104.991
102.197
72.144
69.734
16.319


OCD3
H3C
O
OH
HO
OH
CD3-α-L-Fuc
Normalized Intensity
Chemical Shift (ppm)
4.635
4.630
3.927
3.914
3.901
3.888
3.719
3.714
3.710
3.647
3.640
3.310
1.223
1.210
1.00
1.34
2.48
1.36
3.37

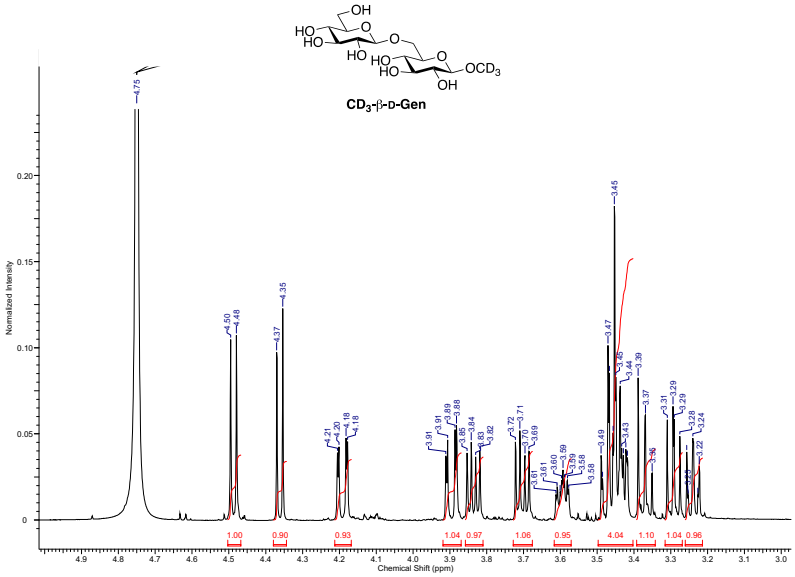

CD_3-β-L-Fuc
Normalized Intensity
Chemical Shift (ppm)

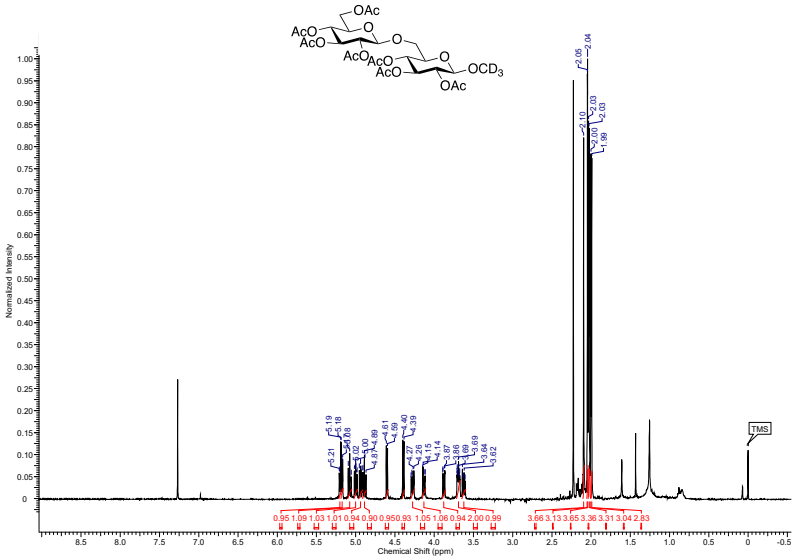

TMS
Normalized Intensity
Chemical Shift (ppm)

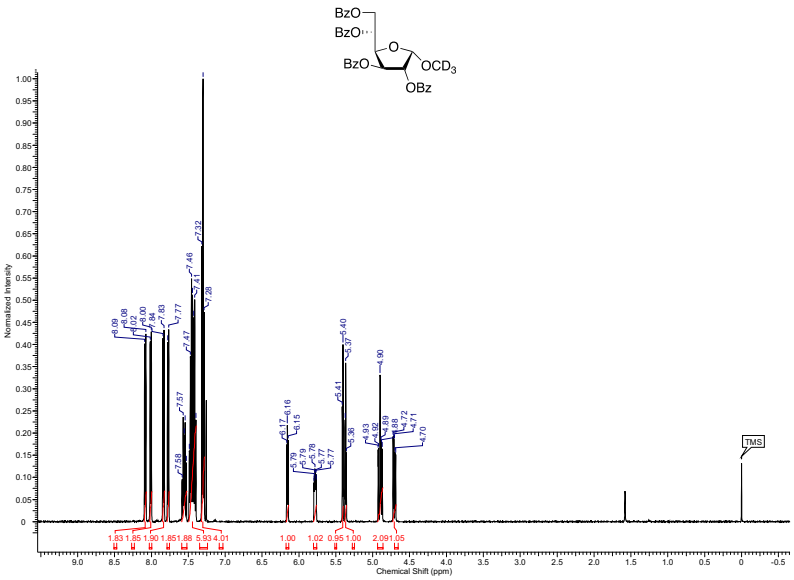

OBz
BzO
BzO
BzO
BzO
BzO
BzO
BzO
OCD3
TMS
Normalized Intensity
Chemical Shift (ppm)

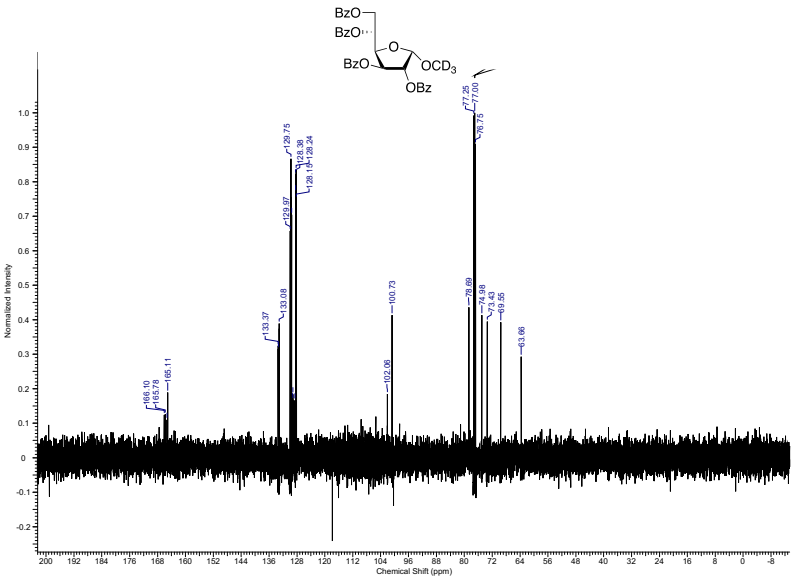

OH
HO
HO
HO
HO
HO
HO
HO
OCD3
Normalized Intensity
Chemical Shift (ppm)

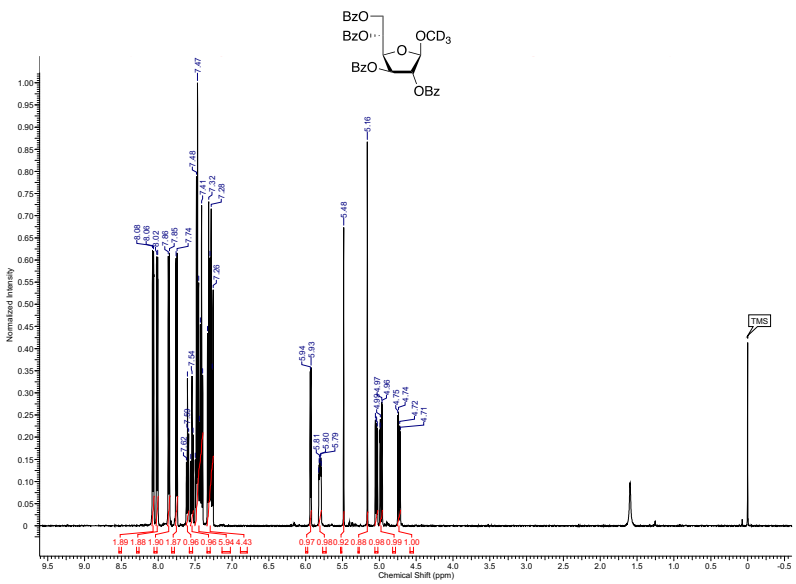
CD_3-α-D-Gen

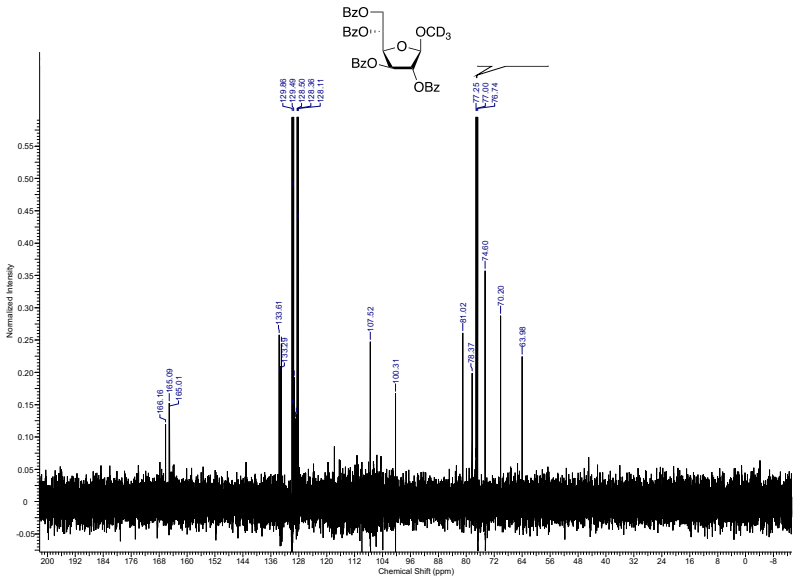

OH
HO
HO
O
HO
O
HO
HO
O
OCD3
OH
CD3-β-D-Gen
Normalized Intensity
Chemical Shift (ppm)

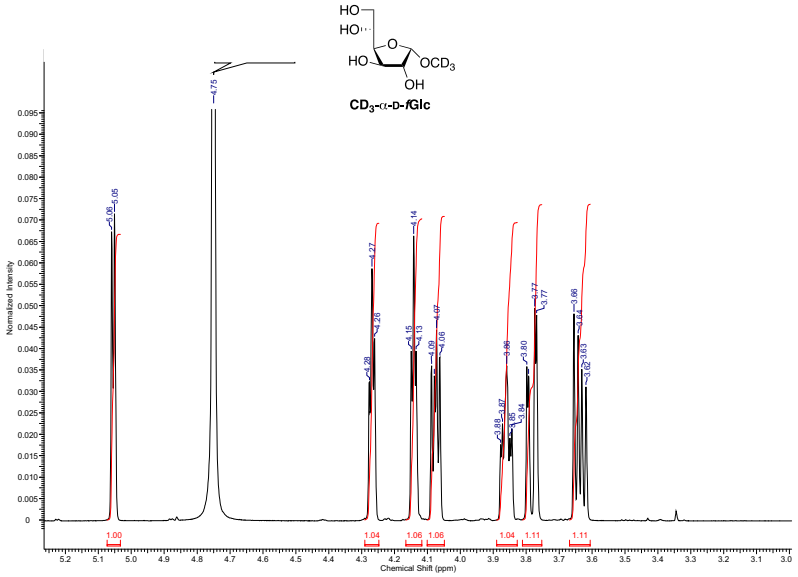

OAc
AcO
AcO
O
AcO
AcO
AcO
O
OCD3
OAc
TMS
Normalized Intensity
Chemical Shift (ppm)

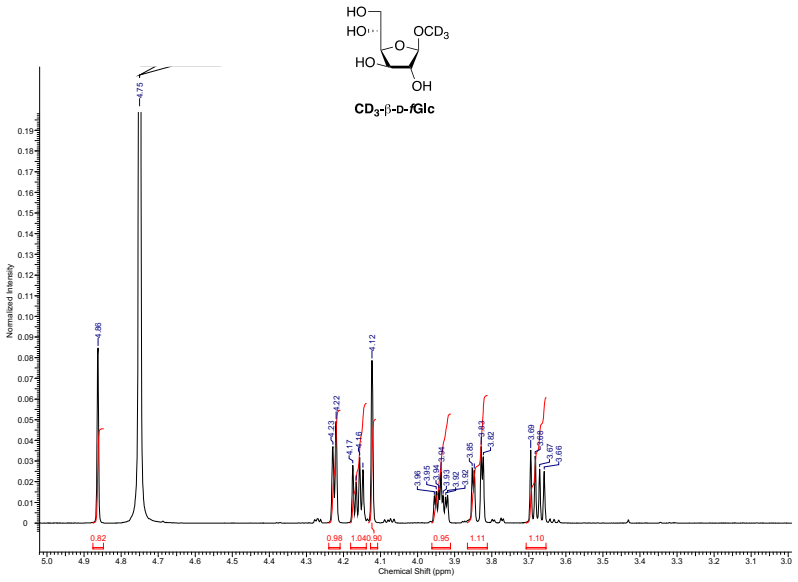

BzO
BzO
O
BzO
OCD3
OBz
TMS
Normalized Intensity
Chemical Shift (ppm)

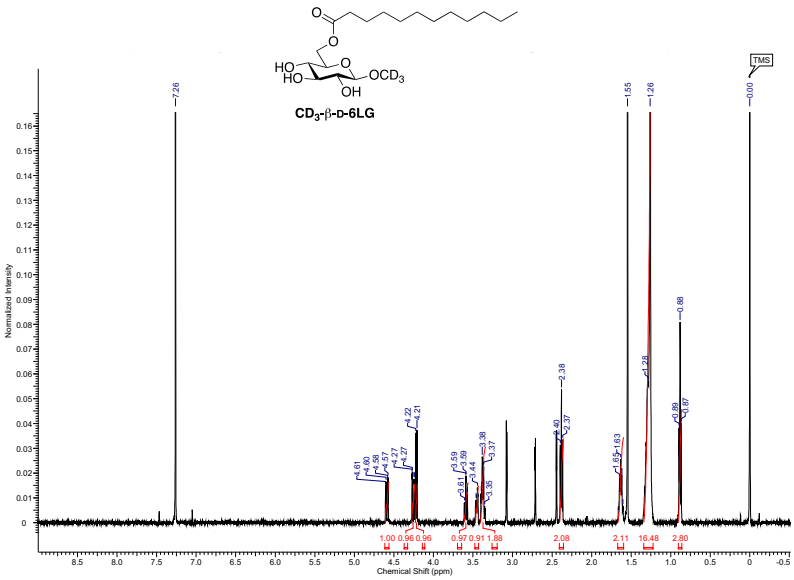

BzO
BzO
O
BzO
OCD3
OBz
Normalized Intensity
Chemical Shift (ppm)

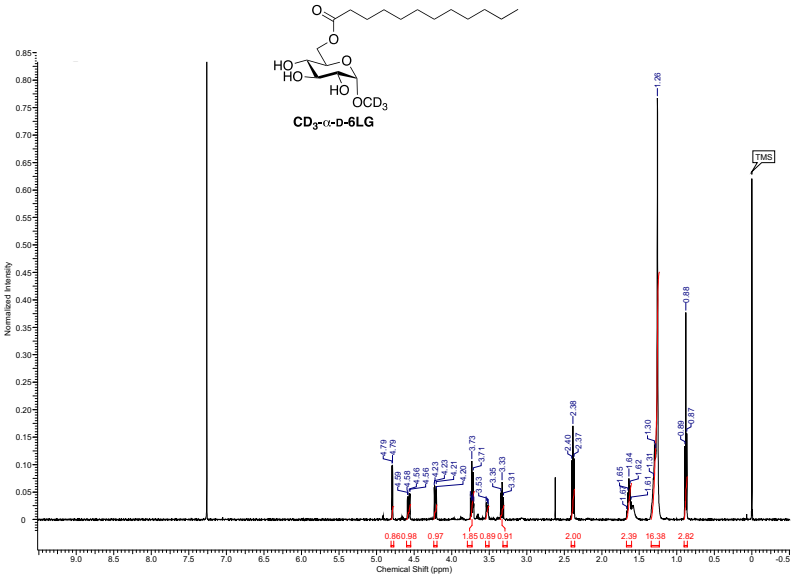

BzO
BzO
OCD3
O
BzO
OBz
TMS
Normalized Intensity
Chemical Shift (ppm)


BzO
BzO
OCD3
O
BzO
OBz
Normalized Intensity
Chemical Shift (ppm)


HO
HO
O
HO
OCD3
OH
CD3-α-D-fGlc
Normalized Intensity
Chemical Shift (ppm)


HO
HO
OCD3
O
HO
OH
CD3-β-D-fGlc
Normalized Intensity
Chemical Shift (ppm)


CD3-β-D-6LG
TMS
7.26
4.61
4.60
4.58
4.57
4.27
4.22
4.21
4.59
4.58
4.44
3.38
3.37
3.35
2.38
2.40
2.37
1.65
1.63
1.55
1.28
1.26
0.89
0.88
0.87
0.00
Normalized Intensity
Chemical Shift (ppm)
1.00 0.96 0.96 0.97 0.91 1.88 2.08 2.11 16.48 2.80


CD3-α-D-6LG
TMS
4.79
4.78
4.59
4.58
4.56
4.55
4.53
4.23
4.21
4.20
3.73
3.71
3.53
3.35
3.33
3.31
2.38
2.40
2.37
1.67
1.65
1.64
1.62
1.61
1.31
1.30
1.26
0.89
0.88
0.87
Normalized Intensity
Chemical Shift (ppm)
0.86 0.98 0.97 1.85 0.89 0.91 2.00 2.39 16.38 2.82

3.7 References

1. Bradley C. Doak, B. Over, F. Giordanetto and J. Kihlberg, *Chemistry & Biology*, 2014, **21**, 1115-1142.
2. V. Poongavanam, B. C. Doak and J. Kihlberg, *Current Opinion in Chemical Biology*, 2018, **44**, 23-29.
3. I. Ohtani, T. Kusumi, Y. Kashman and H. Kakisawa, *Journal of the American Chemical Society*, 1991, **113**, 4092-4096.
4. S. G. Smith and J. M. Goodman, *Journal of the American Chemical Society*, 2010, **132**, 12946-12959.
5. N. Harada and K. Nakanishi, *Journal of the American Chemical Society*, 1969, **91**, 3989-3991.
6. A. Mándi and T. Kurtán, *Natural Product Reports*, 2019, **36**, 889-918.
7. P. L. Polavarapu and E. Santoro, *Natural Product Reports*, 2020, **37**, 1661-1699.
8. J. M. Batista, E. W. Blanch and V. D. Bolzani, *Natural Product Reports*, 2015, **32**, 1280-1302.
9. C. Merten, T. P. Golub and N. M. Kreienborg, *Journal of Organic Chemistry*, 2019, **84**, 8797-8814.
10. E. De Gussem, W. Herrebout, S. Specklin, C. Meyer, J. Cossy and P. Bultinck, *Chem-Eur J*, 2014, **20**, 17385-17394.
11. S. Tanaka, Y. Honmura, S. Uesugi, E. Fukushi, K. Tanaka, H. Maeda, K.-i. Kimura, T. Nehira and M. Hashimoto, *The Journal of Organic Chemistry*, 2017, **82**, 5574-5582.
12. Y. Morishita, H. Zhang, T. Taniguchi, K. Mori and T. Asai, *Organic Letters*, 2019, **21**, 4788-4792.
13. Y. Inokuma, S. Yoshioka, J. Ariyoshi, T. Arai, Y. Hitora, K. Takada, S. Matsunaga, K. Rissanen and M. Fujita, *Nature*, 2013, **495**, 461-466.
14. C. G. Jones, M. W. Martynowycz, J. Hattne, T. J. Fulton, B. M. Stoltz, J. A. Rodriguez, H. M. Nelson and T. Gonen, *ACS Central Science*, 2018, **4**, 1587-1592.
15. A. G. Petrovic, P. L. Polavarapu, J. Drabowicz, P. Łyżwa, M. Mikołajczyk, W. Wieczorek and A. Balińska, *The Journal of Organic Chemistry*, 2008, **73**, 3120-3129.

16. L. G. Felippe, J. M. Batista Jr, D. C. Baldoqui, I. R. Nascimento, M. J. Kato, Y. He, L. A. Nafie and M. Furlan, *Organic & Biomolecular Chemistry*, 2012, **10**, 4208-4214.
17. B. Gordillo-Román, J. Camacho-Ruiz, M. A. Bucio and P. Joseph-Nathan, *Chirality*, 2013, **25**, 939-951.
18. T. Taniguchi, K. Nakano, R. Baba and K. Monde, *Organic Letters*, 2017, **19**, 404-407.
19. D. P. Demarque, D. R. Pinho, N. P. Lopes and C. Merten, *Chirality*, 2018, **30**, 432-438.
20. F. Passareli, A. N. L. Batista, A. J. Cavalheiro, W. A. Herrebout and J. M. Batista, *Physical Chemistry Chemical Physics*, 2016, **18**, 30903-30906.
21. B. Simmen, T. Weymuth and M. Reiher, *The Journal of Physical Chemistry A*, 2012, **116**, 5410-5419.
22. T. Taniguchi and K. Monde, *Journal of the American Chemical Society*, 2012, **134**, 3695-3698.
23. T. Hongen, T. Taniguchi, S. Nomura, J. Kadokawa and K. Monde, *Macromolecules*, 2014, **47**, 5313-5319.
24. T. Taniguchi, D. Manai, M. Shibata, Y. Itabashi and K. Monde, *Journal of the American Chemical Society*, 2015, **137**, 12191-12194.
25. T. Taniguchi, *Bulletin of the Chemical Society of Japan*, 2017, **90**, 1005-1016.
26. T. Taniguchi, M. Z. M. Zubir, N. Harada and K. Monde, *Physical Chemistry Chemical Physics*, 2021, **23**, 27525-27532.
27. D. Crich, A. Banerjee and Q. Yao, *Journal of the American Chemical Society*, 2004, **126**, 14930-14934.
28. T. Taniguchi, K. Monde, N. Miura and S. Nishimura, *Tetrahedron Letters*, 2004, **45**, 8451-8453.
29. D. Kahne, S. Walker, Y. Cheng and D. Van Engen, *Journal of the American Chemical Society*, 1989, **111**, 6881-6882.
30. A. Lubineau and J.-C. Fischer, *Synthetic Communications*, 1991, **21**, 815-818.
31. T. Tamura, S. Shimizu, Y. Sasaki and C. Hirai, *Journal of Japan Oil Chemists' Society*, 1991, **40**, 321-325.
32. A. Smith, P. Nobmann, G. Henehan, P. Bourke and J. Dunne, *Carbohydrate Research*, 2008, **343**, 2557-2566.
33. J. Mutschler, T. Rausis, J.-M. Bourgeois, C. Bastian, D. Zufferey, I. V. Mohrenz and F. Fischer, *Green Chemistry*, 2009, **11**, 1793-1800.

34. A. Esmurziev, N. Simić, E. Sundby, B. H. Hoff, *Magn. Reson. Chem.* 2009, **47**, 449.
35. M. Gutiérrez, T. L. Capson, H. M. Guzmán, J. González, E. Ortega-Barría, E. Quiñoá, R. Riguera, *J. Nat. Prod.* 2006, **69**, 1380.
36. Y. Kondo, *Carbohydr. Res.* 1982, **107**, 303.
37. S. Prabhakar, L. Lemiegre, T. Benvegnu, S. Hotha, V. Ferrieres, L. Legentil, *Carbohydr. Res.* 2016, **433**, 63.
38. S. R. Sanapala, S. S. Kulkarni, *Chem. Eur. J.* 2014, **20**, 3578.
39. V. Gannedi, A. Ali, P. P. Singh, R. A. Vishwakarma, *J. Org. Chem.* 2020, **85**, 7757.
40. D. Crich, S. Sun, *J. Org. Chem.* 1997, **62**, 1198
41. N. Shimada, Y. Nakamura, T. Ochiai, K. Makino, *Org. Lett.* 2019, **21**, 3789.
42. Y. Nishi, T. Tanimoto, Biosci. *Biotechnol. Biochem.* 2009, **73**, 562.
43. A. Khaled, O. Piotrowska, K. Dominiak, C. Augé, *Carbohydr. Res.* 2008, **343**, 167.
44. F. Zhang, W. Zhang, Y. Zhang, D. P. Curran, G. J. Liu, *Org. Chem.* 2009, **74**, 2594.
45. B. W. Skelton, R. V. Stick, D. M. G. Tilbrook, A. H. White, S. J. Williams, *Aust. J. Chem.* 2000, **53**, 389.
46. H. He, X. Zhu, *Org. Lett.* 2014, **16,** 3102.
47. P. -E. Jansson, L. Kenne, I. Kolare, *Carbohydr. Res.* 1994, **257**, 163.
48. J. Ortner, M. Albert, K. Terler, W. Steiner, K. Dax, *Carbohydr. Res.* 2000, **327**, 483.
49. Z. Shi, L. Sun, C. J. Li, *Agric. Food Chem.* 2014, **62,** 3287.

Acknowledgement

First and foremost, in the Name of Allah, The Most Compassionate, Most Merciful. All praise to Allah, The Lord of All Word. The Almighty God has granted me the guidance, strength, patience, and blessings to successfully complete this work.

It is my pleasure to acknowledge my supervisors, Professor Dr Kenji Monde, and Associate Prof Dr Tohru Taniguchi for their complete support, obliging supervision, brilliant suggestions, unlimited tolerance, and compassionate encouragement throughout my study.

Also, sincere thanks go to all lab members and the supporting staff of Faculty of Life Science, Hokkaido University for their help in both academic and other technical affairs that enable me to complete my study.

Above all, I owe my deepest gratitude and appreciation to my beloved wife who stand by me throughout this journey. Also, my parents, family and close friends for their love, concern, and motivational support. None of this would have been possible without all of them. My special thanks to all of you.

Finally, I am grateful for the financial support from MEXT scholarship through International Graduate Program (IGP), Hokkaido University, which made my study in Japan possible.

LIST OF PUBLICATION

1. Papers (related to Doctoral Dissertation)

(1) Mohamad Zarif Mohd Zubir, Nurul Fajry Maulida, Yoshihiro Abe, Yuta Nakamura, Mariam Abdelrasoul, Tohru Taniguchi, Kenji Monde : Deuterium labelling to extract local stereochemical information by VCD spectroscopy in the C–D stretching region: a case study of sugars, ***Organic & Biomolecular Chemistry***, 20, 1067-1072, (2022)

(2) Tohru Taniguchi, Mohamad Zarif Mohd Zubir, Nobuyuki Harada, Kenji Monde : Exploration of chromophores for a VCD couplet in a spectrally transparent infrared region for biomolecules, ***Physical Chemistry Chemical Physic***, 23, 27525-27532, (2021)

2. Papers (Other than the above)

(1) Mahadeva M. M. Swamy, Mohamad Zarif Mohd Zubir, Mutmainah Mutmainah, Setsuko Tsuboi, Yuta Murai, Kenji Monde, Ken-ichi Hirano, Takashi Jin : A near-infrared fluorescent long-chain fatty acid toward optical imaging of cardiac metabolism in living mice ***Analyst***, (2022), Accepted Manuscript

3. Conference (related to Doctoral Dissertation)

(1) Mohamad Zarif Mohd Zubir, Nurul Fajry Maulida, Tohru Taniguchi, Kenji Monde : Exploration of Biomolecularly Transparent IR Region for Structural Identification of Biomacromolecules Using VCD, Molecular Chirality (Nov 2021, Hiroshima, Japan)

(2) Mohamad Zarif Mohd Zubir, Nurul Fajry Maulida, Tohru Taniguchi, Kenji Monde : Exploration of Biomolecularly Transparent IR Region for Structural Identification Using VCD, The 102nd CSJ Annual Meeting, (March 2022, Kobe, Japan