Charles University

Faculty of Science

Study Programme: Physical Chemistry

Ing. Markéta Christou

NMR SPECTROSCOPY STUDIES OF PHOSPHORUS-CONTAINING COMPOUNDS

NMR SPEKTROSKOPICKÉ STUDIE SLOUČENIN OBSAHUJÍCÍCH FOSFOR

Dissertation

Supervisor: RNDr. Eliška Procházková, Ph.D.

Prague, 2024

Declaration

Prohlašuji, že jsem práci sepsala samostatně, všechny použité zdroje a literatura jsou v práci řádně citovány a práce nebo její podstatná část nebyla využita jako závěrečná práce k získání jiného nebo stejného akademického titulu.

V Praze dne ………………………   …………………………………..

                                                                                    Ing. Markéta Christou

Acknowledgements

Contents

List of Abbreviations

APT	attached proton test
ATP	adenosine triphosphate
BINAP	2,2'-bis(diphenylphosphino)-1,1'-binaphthyl
CACO	dimethylarsinic acid sodium salt trihydrate
CD	circular dichroism
COSY	correlated spectroscopy
COVID-19	coronavirus disease 2019
CPCM	conductor-like polarisable continuum model
DFT	density functional theory
DMNB	4,5-dimethoxy-2-nitrobenzyl group
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
ECD	electronic circular dichroism
EDG	electron-donating group
EWG	electron-withdrawing group
FID	free induction decay
GGA	generalised gradient approximation
GIAO	gauge-including atomic orbital
GM	ground minimum
HEPES	2-[4-(2-hydroxyethyl)piperazin-1-yl]ethanesulfonic acid
HIV	human immunodeficiency virus
HMBC	heteronuclear multiple bond correlation
HR-MS	high resolution mass spectrometry
HSQC	heteronuclear single quantum coherence
CLAP-HSQC	clean-antiphase HSQC
CLIP-HSQC	clean-inphase HSQC
IGLO	individual gauge for localised orbitals
IOCB	Institute of Organic Chemistry and Biochemistry of the Czech Academy of Sciences
IPC	isopinocampheol
LDA	local density approximation
LED	light emitting diode
MD	molecular dynamics
MDOC	molecular dynamics with orientational constraints
MP2	Møller-Plesset perturbation theory
MTPA	α-methoxy-α-trifluoromethylphenylacetic acid
NMR	nuclear magnetic resonance
1D NMR	one-dimensional NMR
2D NMR	two-dimensional NMR
NOE	nuclear Overhauser effect

PBLG	poly-γ-benzyl-L-glutamate
PCM	polarisable continuum model
ProTide	prodrug nucleotide
RDC	residual dipolar coupling
RNA	ribonucleic acid
SI	self-immolation
TD DFT	time-dependent density functional theory
UV	ultraviolet
UV/vis	ultraviolet/visible light
VCD	vibrational circular dichroism
XRD	X-ray diffraction

Abstract

Phosphorus is a biogenic element and a key element in modern organic and medicinal chemistry. The non-metallic catalysis based on chiral phosphates is a fast-evolving field similarly to phosphorus-based prodrugs. Phosphorus is also widely used in the industry, e.g. in pesticides or food additives. Many of the compounds named above bear a chiral centre on the phosphorus atom. Thus, it is critical to determine their structure and stereochemistry because it controls their physico-chemical properties. For structural and stereochemical analysis, NMR spectroscopy is a standard method, usually based on ^{1}H, ^{13}C, ^{15}N, ^{19}F or ^{31}P isotopes. ^{31}P NMR spectroscopy can also be ideal for reaction monitoring as ^{31}P spectra are less complicated than ^{1}H, and ^{31}P is more sensitive than ^{13}C nucleus.

In this work, I used ^{31}P NMR spectroscopy to monitor the photo-initiated fragmentation of phosphate-based self-immolative (SI) linkers. The linkers are decomposed in a cascade of chemical reactions via self-immolation based on intramolecular cyclisation resulting in the drug release. I studied a series of newly designed prodrug linkers bearing two cargos. The ^{31}P NMR reaction monitoring with *in situ* irradiation enabled us to observe the reaction course and capture even metastable reaction species in real time. We doubtlessly distinguished which cargo was released preferentially and found structure-activity relationships. This ultimately led to the design of new classes of SI likers suitable for the release of amine-containing drugs, which usually face difficulties in cell-membrane delivery. Using the prodrugs based on the SI process, we successfully released a variety of amine-cargos. I also identified an alternative decomposition pathway of amine-containing cargos with the P-NH-R motif, which could be misinterpreted as the amine-cargo release. However, a careful analysis of the NMR data revealed an alternative decomposition, and I initiated a search for reaction conditions avoiding this undesired decomposition.

During the SI studies, we found no clear trend between the ^{31}P NMR parameters and the molecular geometry, e.g., *J*-couplings vs. the number of chemical bonds between the interacting nuclei, as is usual in ^{1}H NMR spectroscopy. For instance, the two-bond $^2J_{\text{C-P}}$-couplings are often smaller than the three-bond $^3J_{\text{C-P}}$-couplings, etc. Therefore, I decided to investigate how the ^{31}P NMR parameters can contribute to the stereochemical determination.

For the ^{31}P NMR studies, model phosphorus-stereogenic small molecules were designed and prepared as pure diastereoisomers. The ^{31}P NMR parameters were used to study the relative configuration and conformation of model compounds. We complemented the ^{31}P structural analysis with quantum-chemical calculations and a thorough conformational sampling. ^{13}C–^{31}P *J*-coupling analysis unequivocally assigned the relative configuration of rigid molecules, while ^{31}P-based analysis of residual dipolar couplings (RDC) did not provide unambiguous results. This may be caused by insufficient

conformational sampling based on low-energy structures. Thus, we applied molecular docking, generating a new ensemble of conformers in the presence of an alignment medium. This approach improved the results for mildly flexible molecules but did not help for rigid compounds. This indicates that the rigid molecules do not adopt significantly different conformation after interaction with the alignment medium compared to the mildly flexible molecules. For more flexible molecules, we employed a molecular dynamics method with the use of NMR orientational constraints - MDOC. MDOC significantly improved the assignment of the relative configuration, determining even flexible molecules. However, highly flexible molecules with low motion restriction remain a problem which requires attention and further development in NMR methodology.

Abstrakt

Fosfor je biogenní prvek a je klíčový v moderní organické syntéze a medicinální chemii. Asymetrická katalýza reakcí organickými molekulami na bázi chirálních fosfátů (bez přítomnosti kovu) je rychle se vyvíjejícím oborem podobně jako proléčiva na bázi fosforu. Fosfor je také široce využíván v průmyslu, je např. součástí pesticidů nebo potravinářských přísad. Mnoho z výše uvedených sloučenin nese chirální centrum na atomu fosforu. Je tedy důležité určit jejich strukturu a stereochemii, protože ta předurčuje jejich fyzikálně-chemické vlastnosti. Pro stereochemickou analýzu je standardní metodou NMR spektroskopie, obvykle založená na izotopech ^{1}H, ^{13}C, ^{15}N, ^{19}F nebo ^{31}P. ^{31}P NMR spektroskopie může být také efektivně využita pro monitorování reakcí, protože ^{31}P NMR spektra jsou méně komplikovaná než ^{1}H, zároveň je však ^{31}P citlivější než ^{13}C.

V této práci jsem použila ^{31}P NMR spektroskopii ke sledování fragmentace fosfátových self-imolativních (SI) spojek iniciované světlem. Po osvícení fosfátové spojky podléhají intramolekulární cyklizaci následně vedoucí k uvolnění léčiva. Studovala jsem sérii nově navržených spojek proléčiv nesoucích dvě odstupující skupiny. ^{31}P NMR s *in situ* ozařováním nám umožnilo sledovat průběh reakce v reálném čase, a navíc zachytit i metastabilní produkty. Jednoznačně jsme rozlišili, která odstupující skupina byla uvolněna přednostně, a vytvořili jsme strukturně-aktivitní studii. To nakonec vedlo k návrhu nových tříd SI spojek vhodných pro uvolňování léčiv na bázi aminů, které obvykle hůře prostupují přes buněčnou membránu. Pomocí nových fosfátových spojek jsme úspěšně uvolnili řadu aminů. Díky pečlivé analýze NMR dat jsem odhalila alternativní cestu rozkladu spojek obsahujících P-NH-R uskupení, což by mohlo být zaměněno za úspěšné uvolnění aminu. Tyto výsledky vedly k hledání nových reakčních podmínek, za kterých tento nežádoucí přesmyk neproběhl.

Během studií SI jsme nenašli žádný jasný trend mezi parametry ^{31}P NMR a molekulární geometrií, např. *J*-interakce vs. počet chemických vazeb mezi interagujícími jádry, jak je obvyklé v ^{1}H NMR spektroskopii. Například skalární interakce přes dvě vazby ($^2J_{\text{C-P}}$–interakce) jsou často menší než ty přes tři vazby ($^3J_{\text{C-P}}$–interakce) atd. Proto jsem se rozhodla prozkoumat, jak mohou ^{31}P NMR parametry přispět ke stereochemické analýze.

Pro ^{31}P NMR studie byly navrženy modelové malé molekuly se stereogenním centrem na atomu fosforu. Tyto látky byly připraveny jako čisté diastereoizomery. ^{31}P NMR parametry byly použity ke studiu relativní konfigurace a konformace modelových sloučenin. Stereochemickou studii se zapojením ^{31}P parametrů jsme doplnili kvantově-chemickými výpočty a důkladnou konformační analýzou. Analýza na bázi ^{13}C–^{31}P *J*-interakcí jednoznačně určila relativní konfiguraci rigidních molekul, zatímco analýza reziduálních dipolárních interakcí (RDC) se zapojením ^{31}P interakcí neposkytla jednoznačné výsledky. To může být způsobeno nedostatečným popisem konformačních rovnováh s využitím struktur energetického minima. Použili jsme tedy molekulární

dokování a vygenerovali nový soubor konformerů v přítomnosti orientujícího média. Tento přístup zlepšil výsledky pro mírně flexibilní molekuly, ale ne pro ty rigidní. To ukazuje, že rigidní molekuly nezaujímají významně odlišnou konformaci po interakci s orientujícím médiem ve srovnání s mírně flexibilními molekulami. Pro flexibilnější molekuly jsme použili metodu molekulární dynamiky s využitím NMR parametrů jako omezení molekulárního pohybu během simulace, MDOC. MDOC významně zlepšil výsledky analýzy – určil relativní konfiguraci flexibilních molekul. Vysoce flexibilní molekuly s velmi málo omezeným pohybem však zůstávají problémem, který vyžaduje pozornost a další vývoj metodologie NMR.

1. Introduction

1.1 Phosphorus

Phosphorus P is a non-metallic chemical element with atomic number 15 found in the 15th group (with nitrogen) and the 3rd period of the periodic table having electronic configuration $[Ne]3s^2 3p^3$. The three singly-filled orbitals allow phosphorus, similarly to nitrogen, to possess oxidation states of −3 (e.g., PH_3) and +3 (e.g., H_3PO_3), forming a covalent bond with a lone pair of electrons. Unlike nitrogen, phosphorus contains d orbitals which grant him the possibility to form a stable octet configuration with oxidation state +5 (e.g., in H_3PO_4) with five covalent bonds. However, care must be taken with the description of the P=O bond which is usually depicted as a double bond even though it is actually a shorter single bond. Phosphorus is also highly reactive; thus, it does not exist on Earth in a free elemental form.

1.1.1 Naturally Occurring Phosphorus Compounds

Phosphorus is a ubiquitous and biogenic element. In nature, it can be found almost exclusively in a phosphate form with oxidation number +5. Moreover, phosphorus is one of the chemical elements responsible for the origin of life on the primordial Earth.[1-3] Phosphorus plays a crucial role in building blocks of living matter:[4] it is present in nucleic acids, which store genetic information. Phosphorus connects two nucleotides (DNA, RNA) together, stabilises them against hydrolysis and retains the macromolecules within a lipid membrane. Moreover, adenosine triphosphate (ATP, Figure 1A) is responsible for cellular energy management; phospholipids (Figure 1B) are cellular membrane constituents; around 85% of phosphorus in the human body is located in bones and teeth;[5] HPO_4^{2-} acts as an intracellular buffer or ionic carrier, etc.

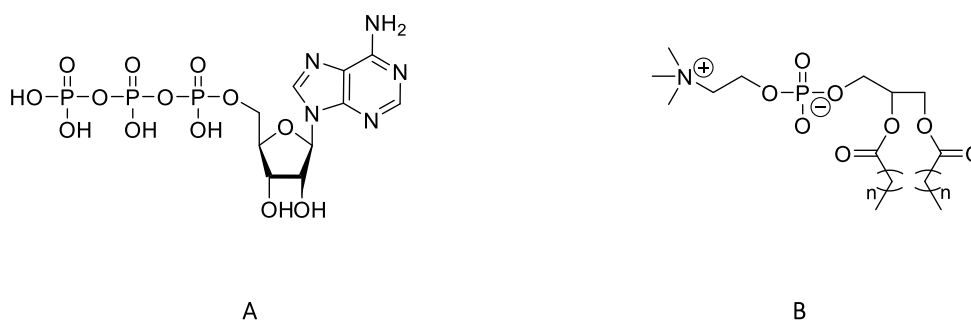A   B

Figure 1: A - Adenosine triphosphate consisting of an adenine base, ribose sugar and three phosphate groups and B - a general phospholipid structure.

1.1.2 Synthetically Prepared Phosphorus Compounds

Phosphorus compounds are utilised in everyday life. Phosphoric acid grants soft drinks their tangy taste; P-based additives increase the efficacy of detergents[6] (eventually being restricted in many countries due to their environmental impact), and phosphorus-containing modifiers in drinking water reduce the risk of lead poisoning.[7] Other phosphorus compounds include food additives[8] (e.g., diphosphate for baking); emulsifiers,[9] flame retardants,[10] plasticisers,[11] pesticides,[12] insecticides,[12, 13] lubricants,[14] solvent extractions agents[15] or compounds part of light-emitting diodes[16] (LEDs). However, such extensive use, especially as fertilizers[17] in agriculture, is unsustainable.

The first commercially available fertiliser was produced in 1841,[18] and the demand has grown ever since.[17] Nevertheless, the phosphorus-rich rock, from which most fertilisers are derived, is a non-renewable material, and its depletion accelerates every year. Moreover, around 95% of phosphorus is lost due to its management inefficiency, which pollutes the environment.[19] Since our current food industry depends entirely on phosphorus fertilisers, the circular economy of phosphorus and its environmental impact represent a challenging task for this century.[7, 17, 20]

From the other side of the spectrum, phosphorus is used in a drug design,[21, 22] especially in the ProTide approach.[23, 24] ProTides are a class of prodrugs which improve the pharmacokinetics and pharmacodynamics of the drug (e.g., the aqueous solubility or cell-membrane penetration of the drug). ProTide (PROdrug + nucleoTIDE) is designed to deliver a nucleotide analogue (monophosphate or monophosphonate) of a drug into cells.[23, 25] The hydroxyl groups of the monophosphate (or monophosphonate) group, masked by an aromatic and an amino acid ester moiety, are enzymatically cleaved-off, releasing a monophosphated (or monophosphonated) drug in the cell. The drug is there further phosphorylated to form a triphosphated (triphosphonated) drug. This second phosphorylation is energetically less demanding than the first phosphorylation of a drug without any phosphate (phosphonate) group attached. The triphosphated drug is then incorporated into DNA, stopping DNA replication. The most famous examples of this group are antiviral drugs tenofovir[26, 27] for HIV treatment (Figure 2), developed by Prof. Antonín Holý at the Institute of Organic Chemistry and Biochemistry of the Czech Academy of Sciences (IOCB), and remdesivir[28] – the only approved medication for COVID-19.

Tenofovir disoproxil fumarate prodrug
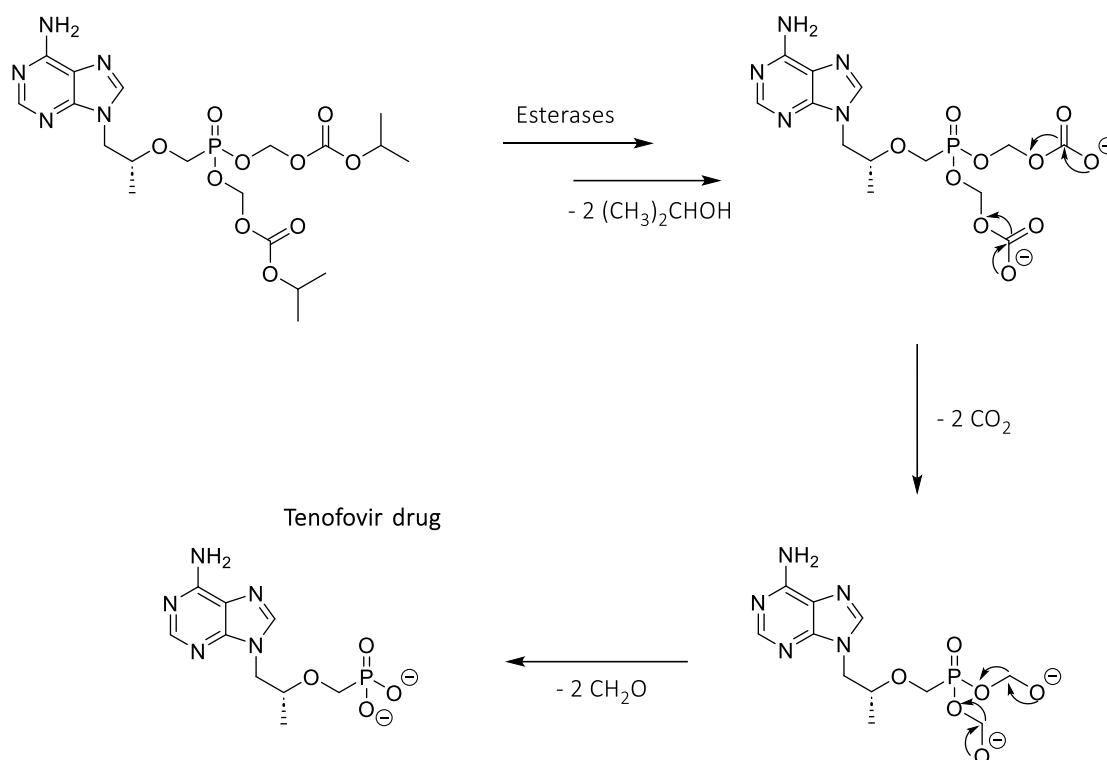
Figure 2: Tenofovir disoproxil fumarate prodrug decomposition pathway.

Modern organic chemistry utilises phosphorus in many ways - as ligands, reaction reagents or catalysts. One of the reasons for this vast usage is phosphorus' dual behaviour – being able to act both as an acid or a base. Acidic behaviour is used in frustrated Lewis pairs[29] (a mixture of Lewis acid and Lewis base, phosphorus is in the form of phosphonium, Figure 3A) in metal-free catalysis, basic behaviour in phosphane superbases[30] (in the form of phosphazenes, Figure 3B). P-based compounds are widely utilised in asymmetric reactions[31] (Noyori reaction with 2,2'-bis(diphenylphosphino)-1,1'-binaphthyl (BINAP) catalyst, Figure 3C), stereoselective catalysis[32, 33] (Josiphos ligands – chiral diphosphines, Figure 3D) and coordination chemistry (Mitsunobu[34, 35] or Wittig[35, 36] reactions, Figures 3E or 3F, respectively). The potential of phosphorus to change its electronic and steric properties makes it a potent element in organic synthesis and green chemistry, both fields constantly growing and evolving.[35, 37]
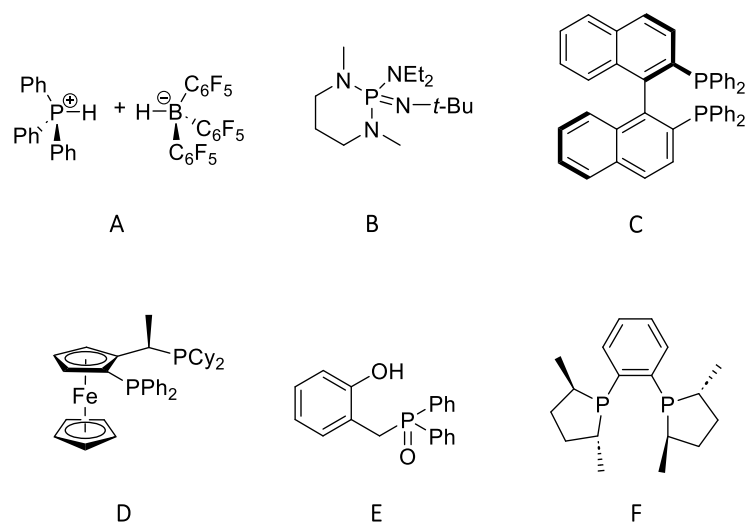
Figure 3: An example of a frustrated Lewis pair (A), phosphane superbase (B), (*R*)-BINAP catalyst (C), Josiphos ligand (D), a catalyst for Mitsunobu reaction (E) and a catalyst for Wittig reaction (F).

In turn, phosphonothioate and phosphonofluoridate groups are also included in chemical warfare agents, such as sarin[38], novichok series[39] or tabun[40] (Figure 4). These compounds possess a stereogenic centre on the phosphorus atom, with the P(-) isomers possessing higher toxicity than their P(+) counterparts.[40] These nerve agents penetrate the dermal membrane and respiratory epithelial, undergo a transformation into their active form (usually by oxidation of different groups, such as thioesters or amides, or hydroxylation of an alkyl group) and inhibit the acetylcholinesterase enzyme responsible for degradation of the neurotransmitter acetylcholine, causing death within minutes.[40]
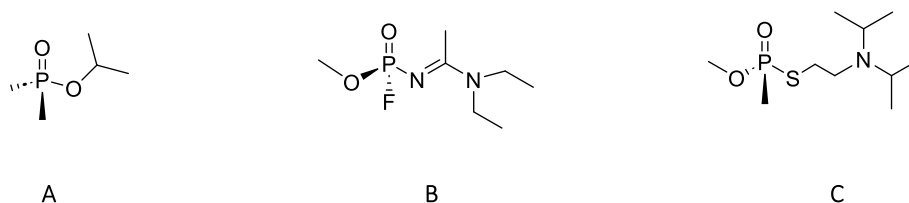
Figure 4: An example of sarin (A), novichok (B) and tabun (C).

Many of the above-mentioned compounds are P-stereogenic (sometimes referred to as P-chirogenic), compounds bearing a chiral centre on the phosphorus atom. However, their faster progression is hindered due to rather difficult synthesis.[41] It is further complicated because, similarly to nitrogen, phosphorus can mutarotate[42] and P-molecules with electron-withdrawing groups are unstable, leading to their decomposition and racemisation already at room temperature.[43]

1.2 Methods for Stereochemical Study

Correctly determining stereochemistry is critical for describing molecular properties and further applications. Absolute (or relative) configuration governs the physico-chemical properties of molecules, such as solubility, dipole moment, acidity/basicity, reactivity, etc. This can have far-range consequences in a pharmacy or material technology as one isomer may possess better selectivity, efficiency, or activity than the other(s). In the presence of an asymmetrical atom (e.g. carbon), the compounds become optically active and can differ in their absolute configuration (enantiomers and diastereoisomers). Moreover, some may coexist in several conformers.

Enantiomers have the same structural formula but differ in the configuration on all stereocentres with respect to each other; thus, they are each other mirror images which cannot be superimposed. Therefore, enantiomers have the same physical and chemical properties, except for the refraction of the plane-polarised light and interaction with other chiral molecules. Importantly, enantiomers can display different biological effects. An infamous example of enantiomers with different biological activity is a Thalidomide drug.[44] Thalidomide was prescribed to pregnant women in the 1960s to help with sleep and anxiety. However, around 10,000 infants whose mothers were prescribed Thalidomide were born with severe defects. It was later found that only the (*R*-) isomer possessed the desired sedative effects, while the (*S*-) isomer was teratogenic. This only stresses the need for an accurate stereochemical investigation of molecules for their further applications.

Diastereoisomers have the same structural formula but differ in the configuration of at least one stereocentre, but not at all of them. Thus, diastereoisomers have different physico-chemical properties. Diastereoisomerism often occurs in carbohydrates which usually bear several stereogenic centres. An example of fundamentally different properties of diastereoisomers is D-mannose with a hydroxyl group in α or β position. While α-D-mannose (with the hydroxyl group in the axial position) tastes sweet, the β-D-mannose (with the hydroxyl group in the equatorial position) is bitter.
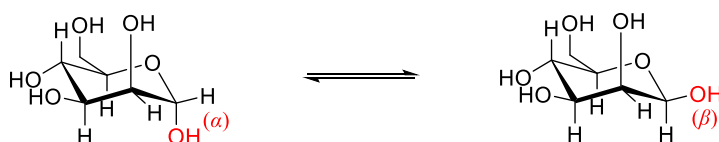
Figure 5: α- (left) and β-D-mannopyranose (right).

Another stereochemical phenomenon is conformation. Conformation is a spatial arrangement of atoms that can interconvert without a bond splitting. Moreover, in some cases, molecules can co-exist in several conformations in equilibrium. If these conformations are stable enough, they can be observed by a suitable stereochemical method and even isolated. Figure 6 shows an example of a conformational isomerism

of cyclopentane,[45] displaying two conformations – a twisted-chair and an envelope with only a small energy difference between them (envelope being more stable only by 0.5 kcal mol^{-1}). Considering the energy difference between both conformers, the population of each conformer in equilibrium may be estimated based on the Boltzmann distribution.

twisted-chair conformation   envelope conformation
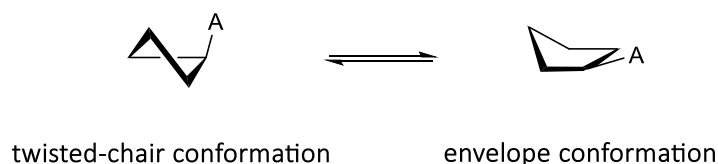
Figure 6: Conformational equilibrium of cyclopentane interconverting between twisted-chair and envelope conformations.

Currently, there are several methods for the study of the stereochemistry (conformation as well as configuration) of the molecule – X-ray diffraction (XRD), chiroptical methods and nuclear magnetic resonance (NMR) spectroscopy.

1.2.1 X-ray Diffraction Analysis

X-ray crystallography is usually a method of choice for determining an absolute configuration if a single crystal is available.[46] The Roentgen beams are diffracted from the crystals, and their subsequently shifted pathways provide information about the electron density, which is, in turn, used to determine the three-dimensional molecular structure. However, the necessity of a suitable single crystal is limiting, especially in highly flexible molecules that are difficult to crystalise.[47] Moreover, the single crystal structure does not necessarily represent the whole sample where a conformational equilibrium may exist, with the XRD analysis showing just one form, usually the more thermodynamically stable conformer.

1.2.2 Chiroptical Methods

The chiroptical methods are based on a different interaction of a chiral molecule with a left- or right-circularly polarised light. Chiroptical methods are non-destructive and, combined with quantum-chemical calculations, can determine the absolute configuration of an optically active molecule.[48] These methods include a wide range of techniques, such as circular dichroism or Raman spectroscopy.

The circular dichroism (CD) method is based on the absorption of circularly polarised light by a chiral molecule. There are two types of CD methods – electronic (ECD) or vibronic (VCD) circular dichroism.[49] ECD technique is used for molecules with an absorption in the ultraviolet region, while the VCD in the infrared region. Moreover, the VCD spectra can be calculated on a density-functional theory (DFT) level with reasonable accuracy, thus further enhancing the stereochemical determination.[50] On the contrary,

for calculations of electronic spectra, time-dependent (TD) DFT must be employed.[51] In turn, Raman optical activity (ROA)[52] is a method detecting the differences in the intensity of Raman scattering of chiral molecules between the right- and left-polarized light. ROA monitors the vibrational optical activity of the molecule. The disadvantages of long measurements and the need for highly concentrated samples are compensated by the possibility of measurement in a broader range of wavelengths compared to VCD. ROA is also suitable for measurements in an aqueous environment.

1.2.3 NMR Spectroscopy

NMR spectroscopy[53, 54] is a versatile technique that offers a wide range of methods for structural investigation in solution and solid state. The main advantage of NMR is that it does not require a single crystal or a thorough analysis complemented by quantum-chemical calculations (but it might be beneficial in some cases – see the following chapters). NMR provides a unique insight into the structural, kinetic, and dynamic data, as described in detail in the next chapter.

1.3 NMR Spectroscopy

NMR spectroscopy[53, 54] measures NMR-active nuclei – nuclei with non-zero spin quantum number I. Spins are quantised which can be described by their magnetic quantum number m. For most common nuclei (e.g., ^{1}H, ^{13}C, ^{15}N, ^{19}F, ^{31}P) with the spin quantum number $I = \frac{1}{2}$, the magnetic quantum number is $m = \frac{1}{2}, -\frac{1}{2}$. Thus, these nuclei can possess two energy levels (states).

In a ground state, spins are oriented chaotically; hence, they have no energy difference. Upon applying an external magnetic field B_0, the spins align in the direction (parallel) or opposite direction (antiparallel) of the magnetisation vector with respect to the applied magnetic field B_0. However, more spins are always aligned in parallel to B_0. This creates an energy difference ΔE between the two spin populations. The energy of one spin is given by equation 1:

$$\Delta E = \gamma h B_0 / 2\pi, \tag{1}$$

where h is the Planck constant and γ is a gyromagnetic ratio specific for each nucleus. Nuclei with higher gyromagnetic ratios can absorb more energy, making them more sensitive than those with lower gyromagnetic values.

Equation 2 shows a magnetic resonance principle upon which a nucleus absorbs energy, and its spin can interconvert between its energy states:

$$\upsilon = \gamma B_0 / 2\pi, \tag{2}$$

where υ is the frequency of oscillating magnetic field B_1. Applying this additional magnetic field B_1, the spins are excited into an energetically higher level. Switching the B_1 field off,

the spins return to their initial state. The relaxation back to the initial state is detected as an NMR signal. Each molecule possesses many specific spin systems, all excited by the applied field B_1. Their frequencies are different from the frequency of B_1; thus, a combination of various frequencies is acquired. They are collected in the form of free induction decay (FID) as an intensity-time dependence. The FID (the time domain) is then mathematically converted into an NMR spectrum (frequency domain) using the Fourier transform.

NMR parameters, such as chemical shift (δ), indirect spin-spin interaction (J-coupling) or dipolar interaction (D-coupling), provide the key structural features.

Chemical shift (a position of a signal in the NMR spectrum) is influenced by the chemical environment of a given nucleus. Each nucleus has its own magnetic field which, in turn, interacts with the magnetic fields of other nuclei. Moreover, it depends on the electron density of neighbouring atoms and functional groups. Regarding stereochemistry, chemical shift allows for discrimination of, e.g., cyclic and acyclic form, *cis* and *trans* isomers, etc.

J-coupling is an indirect interaction of two nuclear spins mediated via chemical bonds. In the NMR spectrum, the value of *J*-coupling can be extracted from NMR signal splitting, which depends on the spin quantum number of the observed isotope and the neighbouring functional groups.

On the contrary, *D*-coupling is a direct interaction of two dipoles through space. In solution NMR, dipolar interaction is averaged to zero due to the Brown motion of the molecule. However, in solid-state NMR, dipolar couplings are prevalent, reaching values in kHz which often hinders the extraction of *J*-couplings (which have values in Hz).

There are several NMR methods which can explore molecular stereochemistry. They could be divided into classical NMR methods, which do not require any special sample preparation, and advanced NMR methods, which often utilise an addition of agents interacting with the analyte molecule, usually destroying the analyte for further use.

1.3.1 Classical NMR Methods

1H NMR Spectroscopy

The most measured nucleus is hydrogen ^{1}H. Isotope ^{1}H has a spin I = ½ and a natural abundance of almost 100%. This, combined with its high gyromagnetic ratio (γ = 26.75 x 10^7 rad T^{-1} s^{-1}),[54] makes it a very sensitive and easily measurable nucleus. The hydrogen ^{1}H is now routinely measured by one-dimensional (1D) NMR spectroscopy (with ^{1}H chemical shifts in the range of 20 ppm). Additionally, two-dimensional (2D) NMR experiments correlating ^{1}H with the ^{1}H distant over three chemical bonds (correlation spectroscopy – COSY experiment) and with other nuclei (e.g. with ^{13}C mentioned below) can provide critical structural information.

From the stereochemical point of view, there is a strong connection between $^3J_{H-H}$-coupling (through three chemical bonds) and the dihedral angle between the interacting ^1H nuclei described by the Karplus equation (Figure 7).[55] The dihedral angles between the interacting nuclei can be estimated from the constructed Karplus curves based on the experimentally obtained $^3J_{H-H}$-couplings. For example, when the $^3J_{H-H}$-coupling is close to zero, it indicates perpendicular ordering (dihedral angle close to 90°). The estimated dihedral angles usually help build a molecular model and propose a possible configuration or conformation.
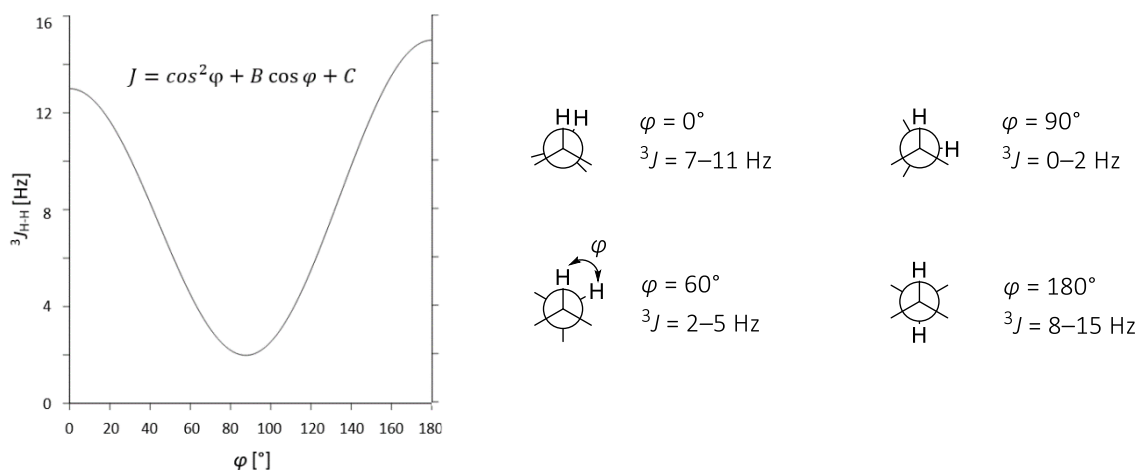
Figure 7: Karplus dependence where φ is the dihedral angle between two vicinal hydrogens and $^3J_{H-H}$ is their experimental J-coupling. The structures on the right show the values of $^3J_{H-H}$-couplings with the corresponding φ.

Complementary to J-couplings, the through-space interactions based on the nuclear Overhauser effect (NOE) can be used for stereochemical analysis. The method of NOE measurements uses a polarisation transfer from an irradiated nucleus to enhance the signal of interacting nuclei through space up to 5 Å, giving the information of a local arrangement.[56] As an example, the NOE enabled to assign the conformation of a thiacalix[4]arene-based compounds.[57] Figure 8 shows a critical NOE contact between H3A (hydrogen atom on A aromatic ring, peak pointing downwards) to OMe, whose signal was enhanced by a polarisation transfer, indicating a partial cone conformation with inverted ring D.
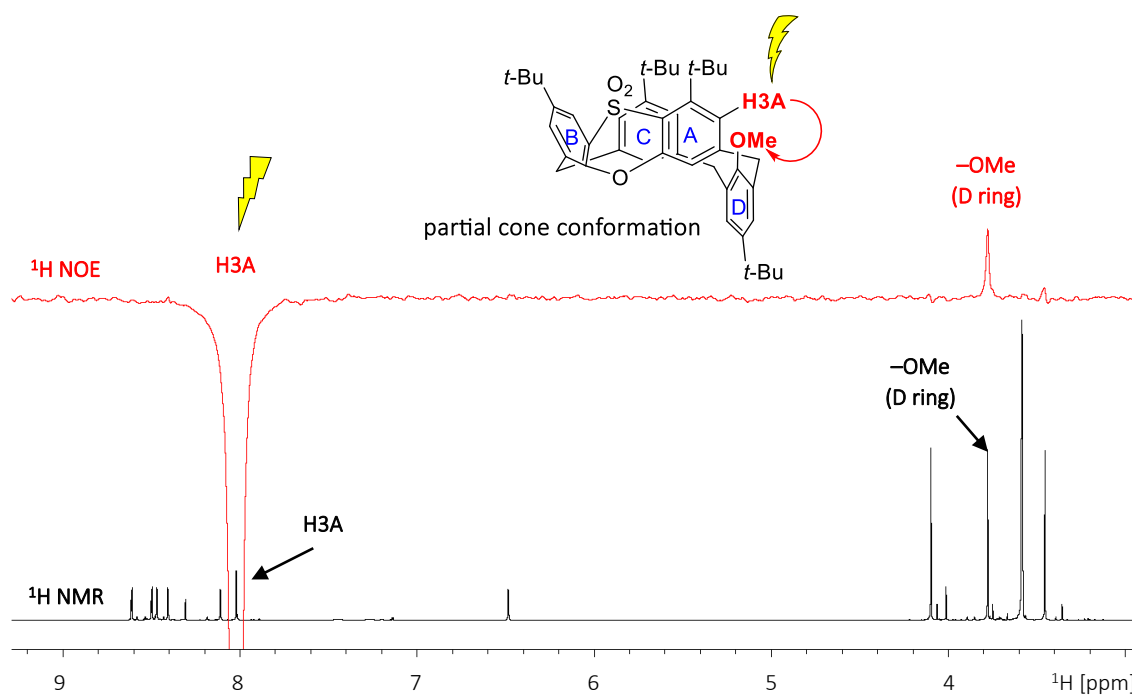
Figure 8: A detail of ^{1}H NMR spectra (in black) with NOE polarisation transfer (red) from H3A to methoxy group (OMe) of the D aromatic ring, indicating their proximity in space which is possible only in a partial cone conformation.

^{1}H NMR experiments are often used for reaction monitoring, given the high sensitivity, fast measurement, and high abundance of ^{1}H in most organic compounds.[58] Chemical transformations, conformational equilibria, or dynamic processes (e.g., sterically hindered rotation) can be studied. Variable temperature NMR spectroscopy is one of the main tools to reveal the above-mentioned phenomena. However, in some cases, e.g., in monitoring a complex reaction course with several intermediates, the ^{1}H NMR spectra may become complicated due to the signal overlaps (usually multiplets in a small chemical shift range). Therefore, other magnetically active nuclei, such as ^{19}F or ^{31}P, can be effectively used.[59]

A special case is monitoring photochemical transformations with reaction intermediates that are not stable enough to be detected by conventional NMR measurements. For this purpose, an NMR with *in situ* irradiation can be crucial for determining the reaction mechanism. This method measures NMR experiments upon continuous UV/vis light irradiation.[60] The system comprises a control unit that manages an LED lamp and an optical fibre leading directly to the NMR tube (Figure 9). This setup allows for recording the NMR spectra upon continuous irradiation in real time and, thus, provides critical structural and kinetic information.

Another advantage of this unique setup is that different wavelengths can irradiate the sample without any sample manipulation (e.g., UV/vis photoswitching reactions).[61-63] Moreover, irradiation is possible at various temperatures, which is beneficial if metastable species are present. Lower temperatures effectively decrease the reaction rate and enable the observation of the metastable species.

| A | B |
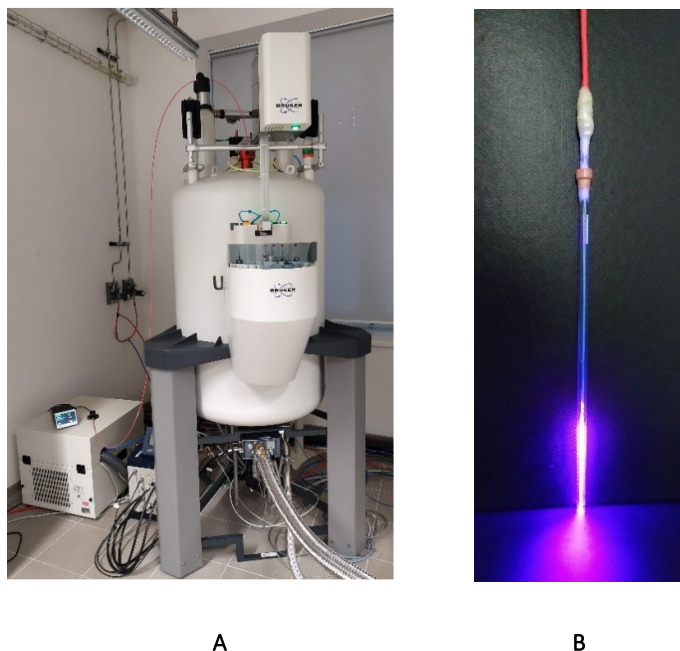
Figure 9: A - complete setup of the NMR with *in situ* irradiation; B – the NMR tube with the directly introduced optical fibre.

¹³C NMR Spectroscopy

Carbon ^{13}C is the second most measured nucleus after ^{1}H. It has a spin quantum number *I* = ½, but natural abundance is only 1.1%, and the gyromagnetic ratio amounts to ¼ of the one of ^{1}H. However, NMR spectrometers equipped with a cryoprobe significantly enhance the sensitivity of ^{13}C NMR. The chemical shifts of ^{13}C are spread over 220 ppm (compared to 20 ppm for ^{1}H), which is beneficial, especially if ^{1}H suffers from signal overlays. ^{13}C is often measured with ^{1}H decoupling where the ^{1}H–^{13}C *J*-couplings are supressed, providing easily-interpretable spectra. This can be applied to reveal interactions with other NMR-active nuclei, such as ^{13}C–^{19}F, especially when the ^{19}F NMR spectra are not affected by ^{13}C–^{19}F interactions due to the low natural abundance of ^{13}C.

Carbon ^{13}C is often utilised in 2D experiments like heteronuclear single quantum coherence (HSQC) or heteronuclear multiple bond correlation (HMBC), which have become standard experiments for structural analysis. The HSQC experiments provide information about directly connected carbon and proton(s). HMBC reveals ^{1}H–^{13}C interaction over two to five chemical bonds. Additionally, without a cryoprobe, HSQC and HMBC allow for indirect detection of ^{13}C NMR with the use of transfer-polarisation from ^{1}H and may reveal signals hidden in 1D, typically quaternary carbon signals or the carbon signals overlapped by the residual solvent signal.

³¹P NMR Spectroscopy

Phosphorus with an atomic number 31 is an NMR-active nucleus with spin ½. Moreover, ³¹P possesses a 100% natural abundance. ³¹P is highly sensitive even to small changes in its chemical environment. ³¹P 1D spectra are usually measured with ¹H decoupling, while the interaction with ¹³C is not visible due to the low natural abundance of ¹³C (1.1%). Therefore, ³¹P spectra are easy-interpretable, displaying singlet signals. This, in combination with a large chemical shift dispersion (up to 1000 ppm), makes it an excellent nucleus for reaction monitoring, given its simplicity (little to no signal overlays compared to ¹H), short measurement times (compared to ¹³C) and sensitivity to its environment. Additionally, phosphorus has a sufficiently large Larmor frequency; thus, it can be easily measured on commonly available NMR probes. ³¹P nucleus also provides interactions with other magnetically active nuclei. E.g., ¹³C–³¹P interactions are easily accessible from the ¹³C NMR signal splitting to doublets (in the presence of one phosphorus atom). Moreover, ³¹P 2D NMR spectra can be measured. But for ³¹P-X correlations (X = ¹³C, ¹⁵N, ¹⁹F, etc.), a special triple resonance probe is needed.

1.3.2 Advanced NMR Methods

As advanced NMR methods, we consider the methods where additional agents interacting with the analyte molecule are used. Advanced NMR methods are usually not routinely measured; however, they can provide valuable information for stereochemical investigations, e.g. absolute or relative configuration. Nevertheless, the analyte usually cannot be recovered after adding the agent to the sample.

Chiral Solvating Agents

Chiral solvating agents[64] (Figure 10) form non-covalent solvating complexes with the molecule studied, inducing signal shifts. These complexes are in a fast exchange with the free solvating agent which is in a large excess compared to the analyte molecule. This method is simple, easy to perform, and does not have kinetic resolution problems. Unfortunately, the agent-substrate interactions are often weak, amounting to only negligible spectral changes.[64]
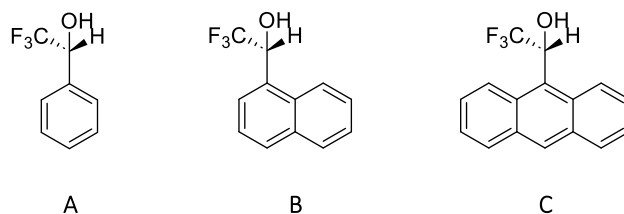
Figure 10: Examples of chiral solvating agents.

Mosher's Method

Mosher's method[65] uses chiral Mosher's acid (α-methoxy-α-trifluoromethylphenylacetic acid, MTPA) to derivatise the analyte – amine or alcohol – and form corresponding diastereoisomers (Figure 11). The determination of the absolute configuration is based on the anisotropic effect of the phenyl group of the MTPA moiety on the A and B substituents of the substrate analysed. Using (*R*)-MTPA, the phenyl group shields the B substituent while the chemical shift of A is unaltered. In turn, employing (*S*)-MTPA has the opposite effect. These shift changes are then applied to assign the absolute configuration of the amine or alcohol studied. The disadvantage of this technique is the narrow substrate scope.
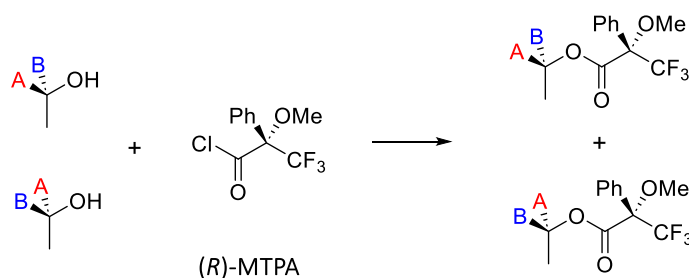
Figure 11: Reaction of (*R*)-MTPA with a mixture of diastereoisomers leading to a different shielding of A and B substituents observable in NMR spectra.

Shift Reagents

The shift reagents are usually based on lanthanide complexes (Figure 12).[66] These six-coordinate reagents form weak complexes or salts with the analytes. Such complexes are in equilibrium with the unbound analyte molecules on the NMR timescale, leading to spectral changes. The scope of analytes is not limited as in the case of Mosher's method; however, these reagents are paramagnetic, leading to a significant line broadening which complicates the spectral analysis. Moreover, this line broadening is proportional to B_0^2, making measuring on lower magnetic fields often preferable.[64]
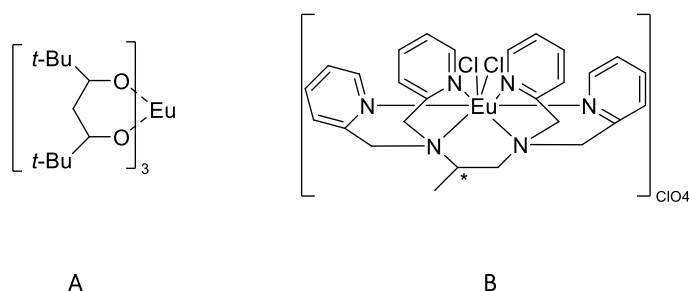
Figure 12: Examples of Eu complexes as shift reagents.[66, 67]

RDC Analysis

RDC analysis is based on dipolar couplings but measured on a solution-state NMR probe. Because RDCs are invisible in solution NMR due to fast molecular reorientations, a weak alignment medium must be added to the sample to partially recover dipolar couplings – now only residual dipolar couplings. Weak alignment media commonly used for RDC analysis can be based on liquid crystals[68-70] or stretched gels.[71, 72]

Each of the two interacting spins, I and S, produces a magnetic field. These fields influence each other and cause a change in the resonances of both spins. The size of this interaction depends on the distance between the interacting spins r_{IS}, the angle θ between the internuclear (interspin) vector and the external magnetic field B_0 (Figure 13).[73] The RDC long-range information is, thus, complementary to *J*-couplings or NOEs.
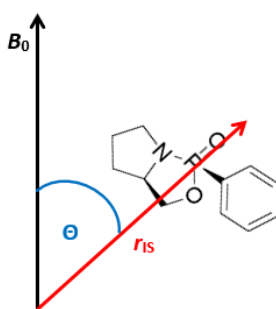
Figure 13: The schematic representation of the RDC dependency on B_0, θ and r_{IS}.

RDC analysis is a powerful, albeit demanding technique due to the challenging sample preparation and the necessity of additional software for back-calculation of the theoretical RDCs. There are three critical steps in RDC analysis which should be mentioned in this work: 1) the sample preparation, 2) the implementation of special NMR experiments, and 3) the use of RDC software for the back-calculation of the theoretical RDCs from the optimised low-energy structures.

1) Sample Preparation

Correct sample preparation is a critical step in RDC analysis. A calculated amount of the alignment medium (usually 6–8% w/w) is dissolved in a corresponding amount of solvent, and a small amount of the analyte is added. Such a mixture becomes highly viscous, and a thorough homogenisation is needed.

To examine the level of homogeneity, a ^{2}H NMR spectroscopy is used either as a 1D ^{2}H experiment[74] or a 2D ^{2}H-image[75] experiment. In liquid crystals, ^{2}H NMR exhibits a quadrupolar splitting Δv_Q[76] due to its electric quadrupole moment which interacts with the gradient of an electric field tensor of the alignment medium. This interaction is observed as a splitting of the deuterated solvent signal (Figure 14A). In turn, the ^{2}H-image experiment uses a gradient to display the homogeneity

within a designated volume of the NMR tube, which is helpful especially as a shimming indicator (Figure 14B).
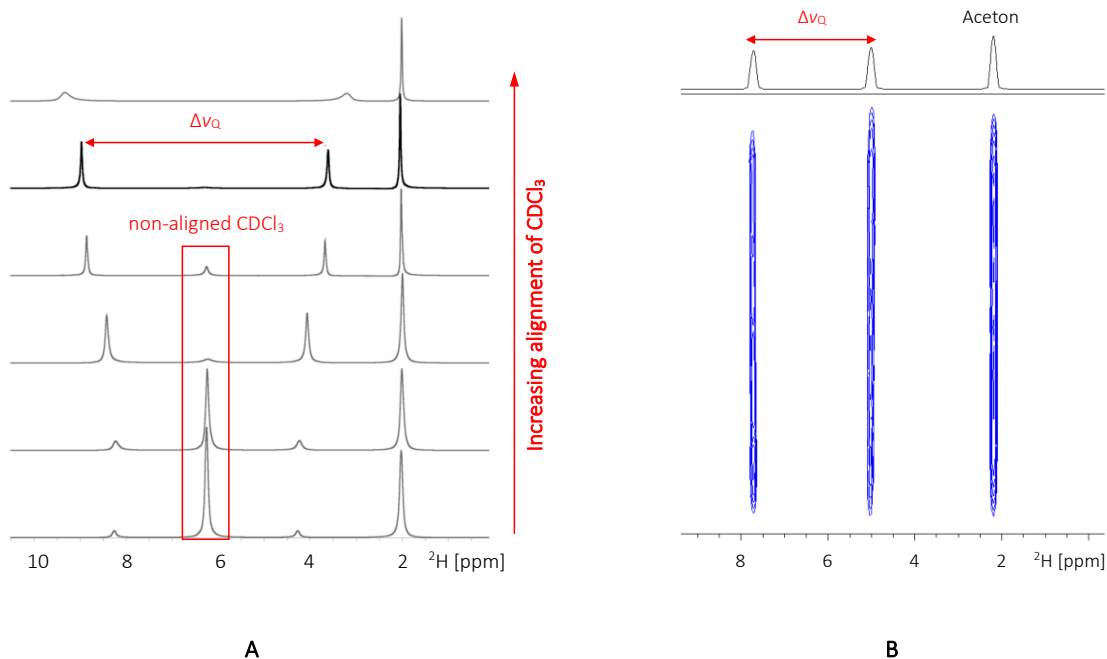
Figure 14: A **-** The alignment of CDCl$_3$ in the alignment medium. The lower ^{2}H spectra show non-aligned CDCl$_3$ which can be caused either by low homogeneity of the sample or an excess of the solvent. The alignment is improving upward, with the black spectrum (5th from the bottom) showing the best sample composition and alignment indicated by narrow signals and Δv_Q (with values around 250–360 Hz signifying the best sample alignment of CDCl$_3$). B - the ^{2}H-image of aligned CDCl$_3$ (in blue) with acetone NMR signal (acetone capillary added for lock).

2) RDC Experiments

Two NMR experiments are necessary to obtain RDCs – an experiment in the isotropic environment (experiment of the analyte with the solvent) and an experiment in the anisotropic environment (the analyte with the solvent and the alignment medium). The final RDC values (D_{X-Y}) are then calculated using equation (3):[77]

$$T_{X-Y} = J_{X-Y} + 2D_{X-Y},$$ 	(3)

with *D* being the residual dipolar coupling of X-Y nuclei, *J* is the *J*-coupling (scalar coupling) obtained from the spectra measured in the isotropic experiment, and *T* is the total coupling obtained from the experiment in the alignment medium (Figure 15).
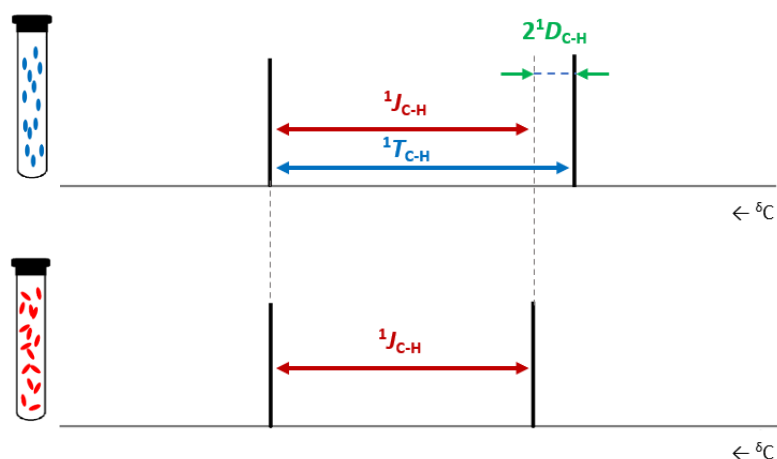
Figure 15: Bottom – ^{13}C NMR experiment in an isotropic environment yielding $^1J_{C\text{-}H}$ coupling (red); top – experiment in the anisotropic environment affording the $^1T_{C\text{-}H}$ coupling (blue), the difference between these two interactions is the residual dipolar coupling $^1D_{C\text{-}H}$ (green).

Experimental ^{1}H–^{13}C RDCs are usually extracted from a ^{1}H-coupled ^{13}C NMR experiment. A more accurate way is to acquire the ^{1}H–^{13}C couplings from an HSQC experiment coupled in the F1 or F2 domain, while the routine HSQC experiment is measured with ^{1}H decoupling (Figure 16A). The HSQC experiment coupled in the F2-domain (Figure 16B) has a relatively short measurement time while achieving high-quality spectra. However, the spectra can be complicated by homonuclear ^{1}H–^{1}H couplings and overlaps accompanied by phase distortions. Several modifications of the F2-coupled HSQC have been made to dispose of these problems (such as CLAP-HSQC, clean-antiphase HSQC, or CLIP-HSQC, clean-inphase HSQC[78, 79]). These problems made F1-coupled HSQC more popular over the years (Figure 16C).[80] An advantage of HSQC coupled in the F1 domain is the high resolution of the two cross-peaks of the corresponding coupling due to the much higher resonance dispersion in the ^{13}C NMR domain (ca 220 ppm compared to 20 ppm in the ^{1}H domain in F2-coupled HSQC).
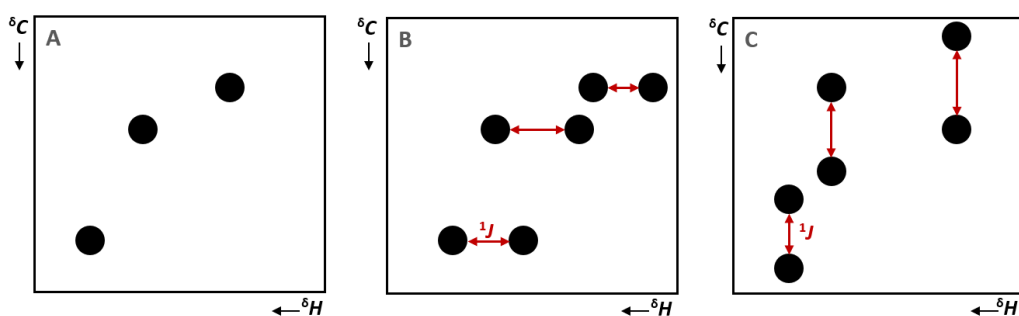
Figure 16: A – standard HSQC experiment, B – F2-coupled HSQC experiment, C – F1-coupled HSQC experiment.

3) RDCs Back-Calculations by the RDC Software

The experimental RDC values are together with optimised structures used as input into a suitable RDC software (Figure 17). This software uses the alignment tensor to back-calculate theoretical RDCs. The alignment tensor associates the molecular frame with the laboratory one and is characterised by five parameters – rhombicity, magnitude, and three axes of molecular alignment.[81] Thus, at least five independent RDC values are needed. While this is easily achieved for rigid molecules, it represents a significant problem for the RDC analysis of flexible molecules. Since θ becomes a variable in time, a single alignment tensor cannot describe the conformations present in the solution. Nowadays, there are several more or less successful approaches trying to solve this problem; the most significant are the multi-alignment-tensor approach,[82] multi-alignment-medium[83] measurements or molecular dynamics-based RDC analysis (MDOC).[84]
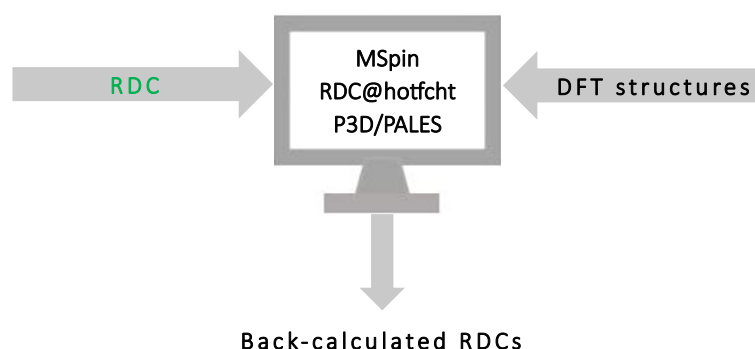
Figure 17: A schematic representation of RDC back-calculations using three different softwares.

The back-calculated theoretical RDC values are then correlated with the experimental RDCs, and quality factor Q^{82} and/or Pearson's correlation factor R designate the quality of the fit. In the case of the Q-factor – the closer to 0, the better the fit; in turn, for R, the closer to 1, the better agreement with the experimental data. Nowadays, there are several software packages for RDC data correlation. For small molecules, MSpin/Stereofitter,[85, 86] RDC@hotfcht[87, 88] and P3D/PALES[89, 90] became the most favourite ones.

MSpin (from MestreLab Company, Spain) was the first commercially available software for RDC analysis. It is based on least-square calculations. MSpin has been recently improved and implemented into MestreNova software as Stereofitter.

RDC@hotfcht was developed in the group of Prof. Christina M. Thiele (Technische Universität Darmstadt, Germany) and works on similar principles to MSpin. However, RDC@hotfcht also considers the experimental error, which can significantly affect the calculations.

P3D/PALES is the newest of the three softwares based on entirely different principle compared to MSpin and RDC@hotchft. It is the first software considering the structure of the alignment medium to calculate the theoretical RDC values. The helical[91] structure of the poly-γ-benzyl-L-glutamate (PBLG, Figure 18) in deuterated chloroform (CDCl$_3$) is embedded within the program as this is the most used system of liquid crystals for the RDC analysis. [70, 92, 93]
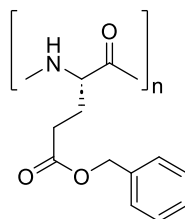
Figure 18: Structure of poly-γ-benzyl-L-glutamate alignment medium.

1.4 Quantum-Chemical Calculations

Stereochemical methods are often combined with quantum-chemical calculations. These calculations complement the experimental data and can be critical for their correct interpretation. For example, the calculated data allow us to determine the conformation and configuration of the molecule studied or may reveal more complex conformational equilibria when the free-energy barrier between two (or more) forms is not high enough to observe distinct forms with the corresponding sets of NMR signals. The first step is the conformational sampling, followed by the geometry optimisation and the calculation of NMR parameters, usually performed using density functional theory (DFT) calculations.

1.4.1   Conformational Sampling

Non-rigid molecules can exist in several conformations, which may interconvert very quickly with respect to the NMR measurement. Only the averaged NMR parameters are then extracted from the NMR spectra. The co-existence of several conformers is observable as a discrepancy between experimental data and the data calculated for each distinct conformer. Therefore, apart from optimisation of the structures, it is also necessary to search the conformational space and find all thermally accessible low-energy conformers. [94, 95]

Conformational sampling has many applications – in structural analysis (such as NMR, chiroptical methods, etc.), pharmacophore modelling[96] or molecular docking,[97] etc. Moreover, different applications require different levels of thoroughness; thus, many automated clustering algorithms have been developed to help in the conformer classification.[98, 99]

The conformational sampling approaches could be divided into systematic and stochastic groups. The systematic methods (such as ConfGen in Schrödinger/Maestro software[100]) require a starting structure to generate conformers by varying all possible torsion angles. These methods are less time-consuming but can result in an excessive number of generated conformations which can overwhelm the generator. The stochastic methods (e.g., ETKDG[101]) employ molecular dynamics[102] to study the evolution of conformation over time; thus, it is a thorough but very time-demanding approach.

The conformations derived from the conformational sampling are then sorted to discard redundant structures and optimised by the below-mentioned methods. The conformations are later used for RDC analysis or a calculation of NMR parameters. These calculated parameters are subsequently weighted by Boltzmann distribution and compared with the experimental data to help in structural analysis.

1.4.2   DFT Calculations

DFT calculations have become an increasingly popular computational method for small organic molecules, usually for geometry optimisation and NMR parameters calculations.[103] The theory uses an electron density matrix instead of a many-particle wavefunction (described by the Hartree-Fock method[104, 105]). The many-particle wavefunction is based on the 3N variable, where N is the number of electrons. However, DFT is a function of only three spatial coordinates, significantly reducing the computational time while attaining a good performance.[106]

The DFT is established on two Hohenberg-Kohn theorems.[107] The first theorem says that the ground energy of a system of interacting particles is a functional of electron density $\rho(\mathbf{r})$. The second Hohenberg-Kohn theorem states that the minimum energy of a system with N-electrons, $E[\rho(\mathbf{r})]$, is equal to the ground energy E_0. DFT is then solved using the Kohn-Sham computational scheme.[108] In this scheme, the exact function of the ground state electron density of the N-electron system is approximated by a system of non-interacting particles. The total energy is then a sum of the Kohn-Sham kinetic energy term, the interaction energy of the interaction with the external magnetic field, Hartree energy and the exchange-correlation energy term. The exchange-correlation energy represents the sums of errors which rise from the approximation used instead of the real interelectronic kinetic and Hartree energy. The exact exchange-correlation functional is unknown; thus, several approaches for its approximations have been developed – the local density (LDA),[109] generalised gradient (GGA),[110] hybrid functionals[111] and meta-GGA approximations.[112] The LDA approximation assumes that the electron density changes only slowly in time. On the other hand, GGA is dependent on density and its gradient. The hybrid functionals include a part of the Hartree-Fock exchange, and the meta-GGAs are the most complex functionals. Currently, the GGA and hybrid functionals (such as B3LYP[111]) are the most commonly used ones in geometry optimisation and calculation of the NMR parameters.

The functionals are combined with basis sets. The basis set is a set of functions used to build molecular orbitals.[113-117,] These functions are one-electron atomic orbitals linearly combined to form the molecular orbitals. In most quantum-chemical calculations, the atomic orbitals are represented by atom-centred Gaussian-type functions. Even though more Gaussian basis functions usually need to be linearly combined, the accuracy and computational cost ratio is still beneficial. The combination of two basis functions forming atomic orbital is called double-zeta; similarly, the combination of three basis functions is triple-zeta, etc.

Several additional parameters complementing the basis set functions are necessary for high accuracy of the quantum-chemical calculations of a studied system. Only the parameters arising from John Pople's group are discussed since they were used in this work:

1) The split valence (designated as a dash in the basis sets entry, e.g. 6-31G) applies only one basis function for each core atomic orbital but uses larger basis functions for the valence atomic orbitals.
2) The polarization function (marked as an asterisk or p/d/f function in a bracket) allows for better charge distribution. Two asterisks indicate the polarization function was also applied to light atoms (hydrogen and helium).
3) The diffuse functions (indicated by '+') are necessary for describing anions, dipole moments, and intra- or intermolecular bonding.

As an example, 6-31+G(d,p) used in our work is a double-zeta split-valence basis set with two polarisation functions and one diffuse function.

Geometry Optimisation

B3LYP complemented with double- or triple-zeta basis sets, usually with one polarization and one diffuse function, in the range of 6-31G to 6-311++G(d,p) Pople's basis sets,[118] are often selected for geometry optimisation, providing good results at a reasonable computational cost.

In some cases, using a GD3 dispersion force correction[119] is beneficial. GD3 is an empirical correction that adjusts the energy of a system and, thus, the final geometry of a molecule under study.

A solvent can also affect the DFT results. There are two main approaches to model solvent effects: 1) implicit solvent models, such as PCM (polarisable continuum model)[120] or CPCM (conductor-like polarisable continuum model)[121] methods, where no specific solvent-analyte interaction is considered, and 2) explicit solvent model,[122] where the solvent molecule is added and its intermolecular interaction with the analyte molecule is introduced. Usually, the implicit model is sufficient for DFT calculations of small organic molecules. The solvent effect in the PCM solvent model is represented

as a cavity surface with a charge distribution. On the other hand, CPCM treats the solvent as a cavity enveloped by a dielectric continuum with a dielectric constant ε.

Calculations of NMR Parameters

The chemical shielding constants can be calculated using various exchange-correlation functional from the GGA group through hybrid functionals to meta-GGAs.[103] However, chemical shielding constant calculations often suffer from the gauge problem which could be described as an unphysical dependence of the calculations on the molecular position in the coordinate frame. This is usually solved using the gauge-including atomic orbital[123] (GIAO) approach, which is embedded in the DFT level nowadays. In GIAO atomic basis sets, the gauge factor (the ratio of fractional change in electrical resistance to the fractional change in length) depends on the magnetic field. This recovers the invariance in relation to the position of gauge origin.

The calculations of ^{31}P chemical shielding constants require a good description of orbitals close to the atomic nucleus. The individual gauge for localised orbitals[124] (IGLO) basis set IGLO-III[125] has been developed to calculate magnetic properties, providing satisfactory results. IGLO basis set can be used for whole molecule or limited only to P atom to save computational time. One of the first significant calculations of ^{31}P chemical shielding constants on small alkylphosphorus compounds[126] (PH_3, $P(CHC_2H_6)_3$, etc.) achieved good correlation with the experimental ^{31}P chemical shifts using B3LYP combined with 6-311++G(2d,2p) basis set. The work of Maryasin and Zipse[127] pointed out that the IGLO-III and 6-311++G(2d,2p) basis sets with MPW1K[128] functional provided ^{31}P NMR chemical shift in good agreement with experimental data, while any lower basis set did not yield such results. They also stated the need for Boltzmann averaging of chemical shifts for larger or more flexible molecules. Finally, for calculating the ^{31}P chemical shift in solution, the addition of solvent effect (the PCM method) proved necessary. Furthermore, Latypov[129] proved that Hartree-Fock, DFT or Møller-Plesset MP2 perturbation theory, combined with any of Pople's basis sets are enough to provide accurate ^{31}P NMR chemical shifts at a reasonable computational cost.

The literature concerning the calculation of ^{31}P NMR spin-spin (*J*) coupling constants, especially for small molecules, is much less common than the one of ^{31}P NMR chemical shifts and is usually focused on the ^{1}H–^{31}P and ^{13}C–^{31}P *J*-couplings.[130] The quantum-chemical calculations of ^{31}P *J*-couplings are generally treated similarly to those of ^{31}P chemical shielding constants. Combining hybrid functional B3LYP with Pople's (6-311+G(d,p)), GIAO or IGLO basis sets are usually used.[118] The solvent effects must also be considered, albeit no general approach is recommended.

2. Aims of the Work

1) Reaction monitoring and structure determination of novel phosphate-based linkers designed for drug release via self-immolation: sequential release of two cargos monitored by ^{31}P NMR spectroscopy

2) Structure-activity relationship studies of novel phosphate-based self-immolative linkers designed for efficient release of amine-containing cargos

3) Investigation of ^{31}P NMR parameters on stereochemical analysis of model phosphorus-containing compounds

4) Application of molecular docking approach for more realistic conformational sampling used for RDC analysis

5) Implementation of ^{31}P NMR parameters to molecular dynamics with NMR orientational constraints

3. Results and Discussion

3.1 Phosphate-Based Self-Immolative Linkers for Tunable Double Cargo Release (Paper I)

Self-immolation (SI) is a triggered irreversible fragmentation of a molecule.[131] Such fragmentation can be used to design self-degrading materials[132] or drug delivery systems.[133] The molecule (linker) used in SI comprises three parts – a trigger (grey), a spacer (orange) and a cargo (purple) (Figure 19). The fragmentation can be activated enzymatically, chemically or photochemically (by light). Upon activation, the trigger group is cleaved off, followed by a spontaneous intramolecular cyclization of the spacer, resulting in the cargo release.
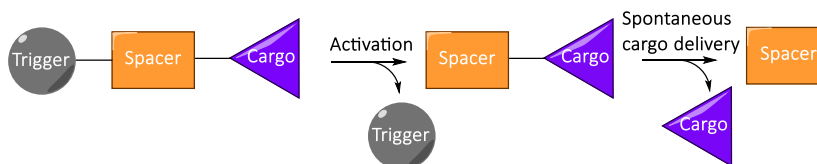
Figure 19: A schematic representation of carbon-based self-immolative linkers.

This work is focused on SI linkers releasing two cargos using a phosphorus-based spacer instead of a traditionally used carbon one.[134] Phosphorus has a higher valency than carbon and, thus, enables us to attach a second cargo (Figure 20). The second cargo release extends the potential scope of applications (e.g. cancer treatment with complementary drugs[135]). Additionally, the phosphorus atom allows for the reaction monitoring by ^{31}P NMR spectroscopy. As a trigger moiety, we used a photoactive 4,5-dimethoxy-2-nitrobenzyl (DMNB) group. This group guarantees the selectivity of the reaction and suppresses any side products. Moreover, unlike enzymes, light-triggering does not struggle with substrate specificity. After activation, the SI process was monitored by ^{31}P NMR spectroscopy with *in situ* irradiation, providing both structural and kinetic information in real time. The compounds studied in this work were prepared by Dr Petr Šimon at the Faculty of Science, Charles University.
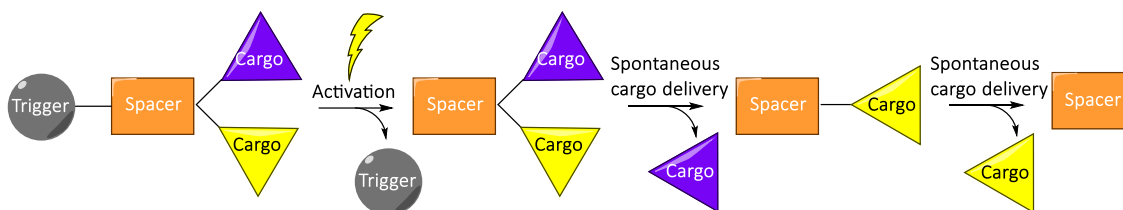
Figure 20: A schematic representation of phosphorus-based self-immolative linkers activated by light studied in this work.

3.1.1 Proof of Concept

We designed model P-chirogenic SI linkers **1** and **2** bearing two phenolic cargos and lactate as an SI spacer (Figure 21). Compound **1** was substituted with an electron-withdrawing (EWG) fluoride in the phenyl *p*-position, while compound **2** bore an electron-donating (EDG) methyl group. Moreover, the compounds possess two stereogenic centres – one on the phosphorus atom and the second on the lactate spacer, giving rise to two diastereoisomers of each **1** and **2**. The compounds were then dissolved in a mixture of cacodylate buffer (dimethylarsinic acid sodium salt, CACO) and deuterated dimethyl sulfoxide (DMSO-d_6; CACO/DMSO-d_6 (1:1, *v/v*)), which was crucial to fix the pH of the solution to 7.4.
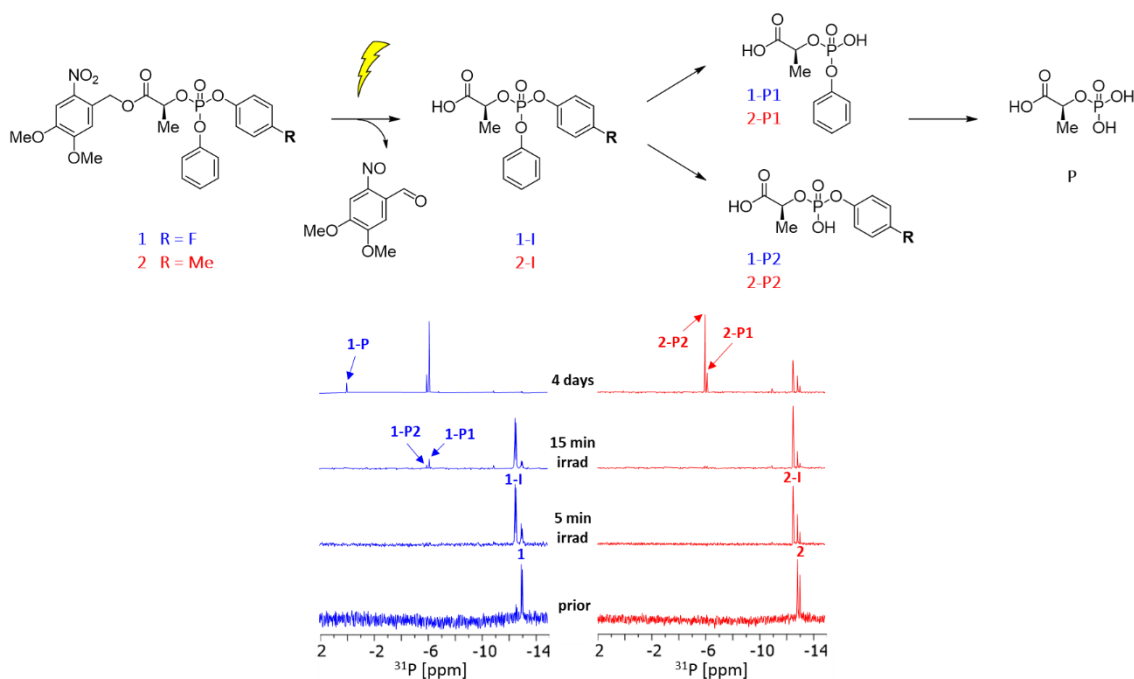
Figure 21: The SI of compounds **1** and **2** of 5 mM solution of **1** (blue) and **2** (red) in 50% CACO/DMSO-d_6 monitored by ^{31}P NMR spectroscopy before and after UV light irradiation (365 nm) at room temperature.

The structures of the two intermediates (**1-I** and **2-I**) and the two key products indicating which cargo was released preferentially (**P1** and **P2**) were determined *in situ* by combining ^{13}C and ^{31}P NMR parameters. We collected the critical ^{13}C–^{31}P interactions from the ^{13}C signal splitting. Additionally, the ^{13}C–^{19}F interactions enabled us to identify which phenyl moiety was released first.

Compound **1** afforded two ^{31}P NMR signals at $\delta_P = -13.02$ and -13.10 ppm (one for each diastereoisomer). The photoactive DMNB group cleaved off after UV light irradiation, forming an intermediate **1-I** ($\delta_P = -12.55$ and -12.64 ppm) in 15 minutes (Figure 21). This activated intermediate **1-I** cyclised followed by a cargo release and finally formed mono-phenyl products **1-P1** ($\delta_P = -6.17$ ppm) and **1-P2** ($\delta_P = -5.98$ ppm) overnight. Product **1-P1** was distinguished from **1-P2** by the missing ^{13}C–^{31}P interactions of the cleaved *p*-F-phenyl group. The more acidic *p*-F-phenyl group (pK_a 9.95[136]) was released preferentially

24

(**1-P1**), while the phenyl substituent (with higher pK_a 9.98[137]) remained mostly untouched (**1-P2**) – only traces found in four days. After four days in the darkness, we could trace a small amount of the hydrolysed final product **P** (δ_P = 0.06 ppm, Figure 21).

Compound **2** (δ_P = −12.92 and −13.09 ppm) provided the intermediate structure **2-I** (δ_P = −12.58 and −12.91 ppm) in 5 minutes. However, the cyclisation process was much slower than in the case of **1**, and the intermediate remained a major component even after 24 hours. After four days, we observed both products **2-P1** and **2-P2** (δ_P = −6.19 ppm and δ_P = −6.02 ppm, respectively), with **2-P2** in the majority, releasing the more acidic unsubstituted phenyl group preferentially.

These data show that although both linkers afforded the same product **P1** (**1-P1** and **2-P1**), the release rates differed significantly (15 minutes for **1** and 24 hours for **2**). This indicates that the nature of the cargo influences which cargo is released preferentially.

3.1.2   Cargo Optimisation

Encouraged by the results mentioned above, we studied the pK_a effect of the cargo on the cargo release. For this, we designed a series of compounds bearing lactate spacer and only one cargo. Compounds **3–7** differed in a substituent in the *para* position (Figure 22).

3	R = NO$_2$	(pK_a = 7.15)
4	R = F	(pK_a = 9.95)
5	R = H	(pK_a = 9.98)
6	R = Me	(pK_a = 10.14)
7	R = OMe	(pK_a = 10.55)
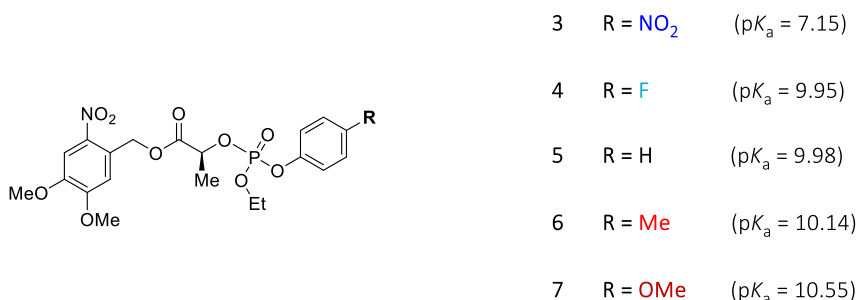
Figure 22: Structures of model compounds **3–7** and the pK_a of the corresponding substituted phenol group.

Linker **3** (δ_P = −8.15 and −8.18 ppm, pK_a = 7.15[137]) offered product **P2** (this product is the same for all compounds **3–7** – a final product after the phenol release) already after 15 minutes of irradiation (δ_P = −1.05 ppm, Figure 23). The SI process was fast, and we did not detect any traces of intermediate **3-I**. The SI of **4** (δ_P = −7.03 and −7.31 ppm, pK_a = 9.95) was slower, offering the intermediate **4-I** after 5 minutes (δ_P = −7.04 and −7.38 ppm) with traces of **P2**. In contrast, the derivative **5** with unsubstituted phenyl ring (δ_P = −7.61 and −7.75 ppm, pK_a = 9.98) provided the product **P2** overnight while still retaining a high concentration of intermediate **5-I** (δ_P = −7.14 and −7.53 ppm). In this case, the intramolecular cyclization is a rate-limiting step in the reaction course. Compound **6** (δ_P = −7.38 and −7.62 ppm, pK_a = 10.14[137]), bearing the electron-donating *p*-Me, offered minor amounts of product **P2** after four days, while compound **7** (δ_P = −6.47 and −6.67 ppm, pK_a = 10.55[138]), substituted with -OMe group, already released cargo after 15 minutes of irradiation. Apart from compound **7**, the release rates matched the pK_a values

25

of the substituted phenyl groups with the trend: **6** (*p*-Me)<**5** (*p*-H)<**4** (*p*-F)<**7** (*p*-OMe)<**3** (*p*-NO$_2$). The unexpected release rate of *p*-OMe derivative **7** might be explained by the resonance effect of a free electron pair of oxygen in the *p*-OMe group, which accelerates the SI.
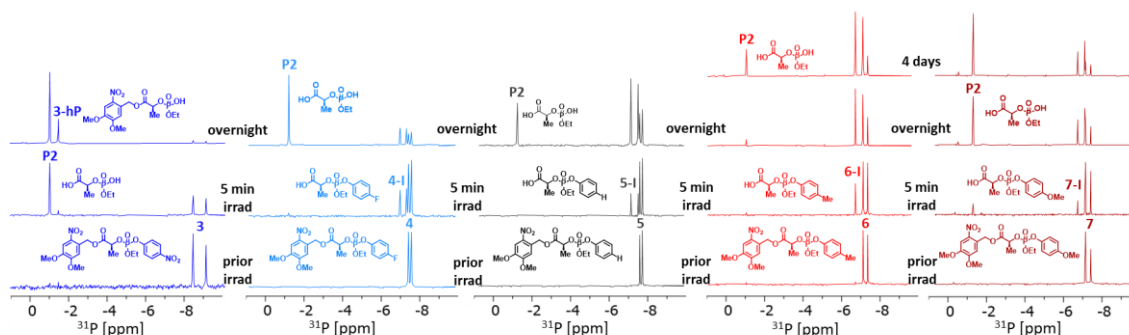
Figure 23: A series of ^{31}P NMR spectra of compounds **3–7** (5 mM, 50% CACO/DMSO-*d$_6$*) recorded upon UV light irradiation (365 nm) at room temperature.

3.1.3   Spacer Optimisation

To increase the release rate, we modified the spacer structure responsible for the SI process. We followed the Thorpe-Ingold effect[139] and designed two additional spacers with sterically demanding substituents in the α position. The α-methyl group from the above-studied lactate (**5**) was modified to *i*Pr (**8**, δ_P = −7.09 and −7.27 ppm) and di-Me (**9**, δ_P = −11.00 ppm) as shown in Figure 24.
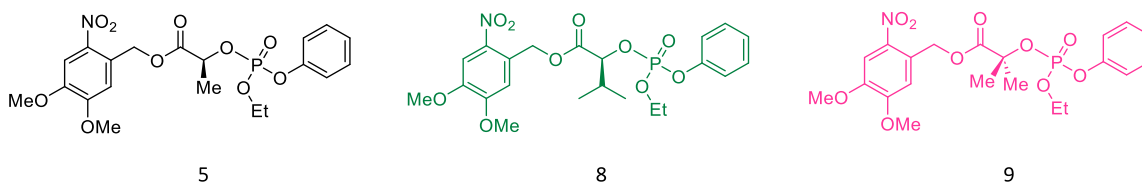
Figure 24: Structures of model phosphate-based SI linkers **5**, **8** and **9** varying in SI spacer.

We did not observe a significant difference in phenol release from **5** and **8**, as similar amounts of the final product **P2** were obtained overnight (Figure 25). These data contradict the Thorpe-Ingold effect, showing that the higher steric demand in **8** does not markedly affect the cargo release rate. In contrast, the α-hydroxyisobutyrate linker **9** with two methyl groups in the α position, provided a high amount of **9-P2** (δ_P = −4.22 ppm) already after 15 minutes of irradiation. The concentration of intermediates **5-I** and **8-I** (δ_P = −6.63 and −7.08 ppm, respectively) grew from 40 to 70% (in 20–60 minutes), while **9-I** (δ_P = −10.85 ppm) remained at 20%. This shows that **9-I** cyclises faster than **5-I** and **8-I** and, thus, accelerates the cargo release significantly.

26
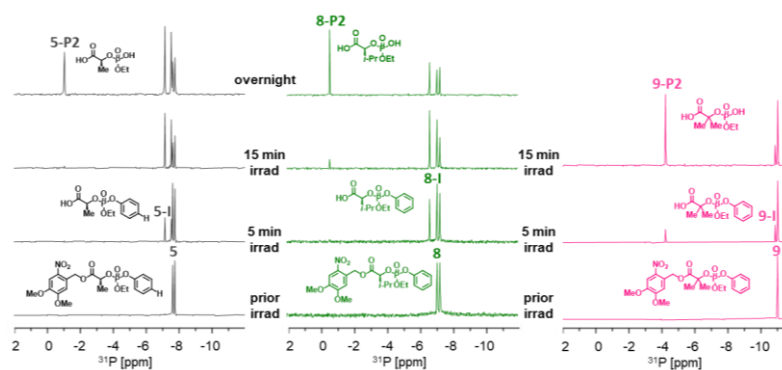
Figure 25: ^{31}P NMR spectra of linkers **5**, **8**, and **9** (5 mM, 50% CACO/DMSO-d_6) recorded upon UV light irradiation (365 nm) at room temperature.

3.1.4 Double-Cargo Linkers with a Tuneable Release Rate

Based on the structure-activity relationship results, we designed new SI linkers **10** (δ_P = −12.45 and −12.49 ppm), **11** (δ_P = −12.38 and −12.47 ppm), **12** (δ_P = −16.50 ppm), **13** (δ_P = −16.49 ppm, Figure 26) for a sequential double-cargo release on model compounds.

10	R_1 = iPr	R_2 = H	R = F
11	R_1 = iPr	R_2 = H	R = Me
12	R_1 = Me	R_2 = Me	R = F
13	R_1 = Me	R_2 = Me	R = Me
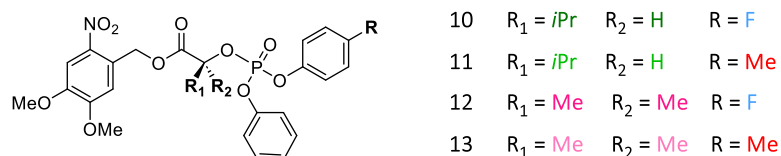
Figure 26: Structures of model compounds **10–13**.

The different spacers affected the SI significantly. The self-immolation with lactate linker proceeded in 24 hours in **1** and **2**, in 2 hours with the α-hydroxyisovalerate spacers **10** and **11** and in 15 minutes with the α-hydroxybutyrate analogues **12** and **13** (Figure 27). This corresponds to the trend found in linkers **5**, **8** and **9**. Moreover, the relative concentrations of **12-I** (δ_P = −16.41 ppm) and **13-I** (δ_P = −16.42 ppm) were approximately 15 and 35 %, respectively, indicating faster SI and subsequent cargo delivery of p-F linker **12**. In general, linkers **1**, **10**, and **12** with electron-withdrawing cargo provided significantly higher amounts of **P1** and/or **P2** than their electron-donating analogues **2**, **11**, and **13**, respectively (Figures 21 and 27), as apparent from the 15th minute of irradiation. The double-cargo linkers **10–13** displayed the desired properties, covering a wide range of cargo release rates – from minutes to days – and enabling sequential cargo release.

27
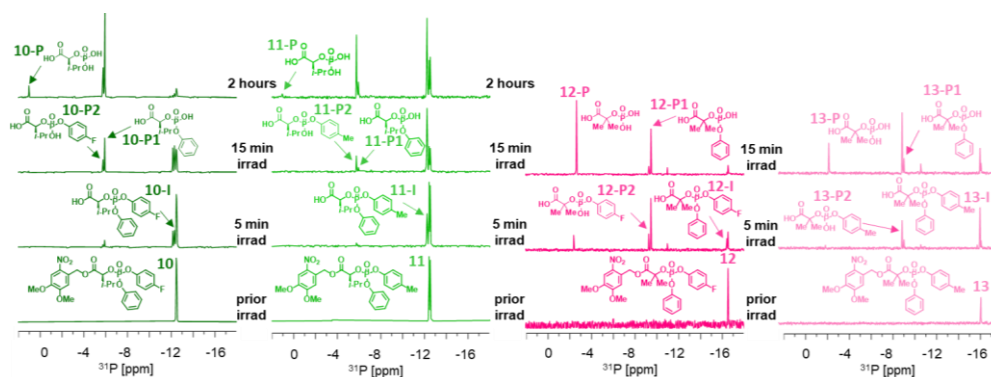
Figure 27: ^{31}P NMR spectra of compounds **10**–**13** (5 mM, 25% (**10**, **11**) or 50% (**12**, **13**) CACO/DMSO-d_6) recorded upon UV light irradiation (365 nm) at room temperature. The α-hydroxyisovalerate analogues **10** and **11** were measured in 25% CACO/DMSO-d_6 due to their low solubility in the 50% solvent system.

3.1.5   Conclusion

In this work, ^{31}P NMR spectroscopy proved essential for studying the structure and properties of phosphate-based SI linkers. The phenol-based cargo is released faster with increasing pK_a of the corresponding phenol. The nature of the second cargo directs the release rate. Following the Thorpe-Ingold effect, sterically demanding SI spacers accelerate the cargo release significantly. Results from this structure-activity relationship study open a new door for a rational design of new tailor-made linkers offering a tuneable double-cargo release option. ^{31}P NMR combined with ^{13}C NMR provided the essential connectivity information. This led us to successfully determine all formed reaction species *in situ* (in reaction mixture without further separation), increasing the screening efficacy.

3.2 Phosphate-Based Self-Immolative Linkers for the Delivery of Amine-Containing Drugs (Paper II)

Amine-containing drugs are part of anti-inflammatory,[140] anticancer,[141] antimicrobial,[142] and pain-relieving pharmaceuticals.[143] Nevertheless, they often display poor aqueous solubility and low membrane permeability[144] under physiological conditions. This may be overcome by the prodrug strategy.[145] Moreover, the self-immolation approach could work well when designing a suitable SI linker.

As we have already shown in section 3.1, the drug release rate can be tuned by a spacer. Additionally, the cargo release is strongly influenced by the nature of the second cargo attached to the phosphorus core. Therefore, we tested different spacers in combination with two types of second cargo to find a suitable system for the release of amine-containing drugs.

3.2.1 Ethylene Glycol Phosphate-Based Linkers

Based on the previous study,[146] we started with an ethylene glycol spacer to design novel SI linkers releasing amine-containing drugs. The ethylene glycol spacer is the only spacer providing stable cyclic intermediate **cyc-I** which can be detected by ^{31}P NMR spectroscopy (Figure 28). Thus, asserting the cargo release through self-immolation.[146]
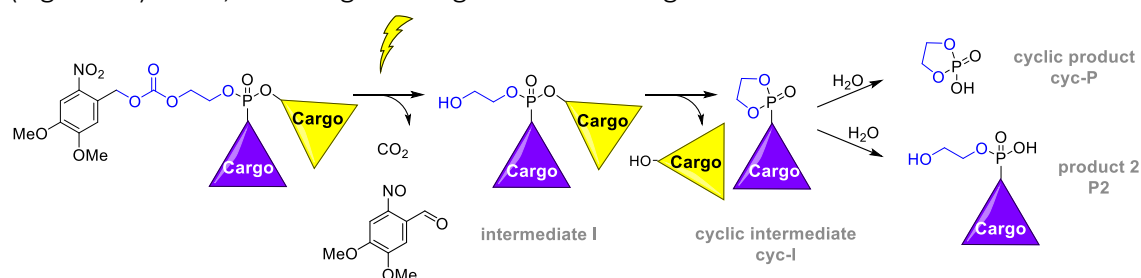
Figure 28: The SI pathway of ethylene glycol phosphate-based linkers.

In this work, we prepared a series of phosphate-based SI linkers bearing amine cargo with phenyl (or ethyl) as the second cargo. We designed a series of model glycol-based linkers **14–19** (Figure 29) carrying phenethylamine, piperidine, and aniline as representatives of primary, secondary, and aromatic amines, respectively. The tested compounds were prepared by Dr Mateja Đud from the Faculty of Science at Charles University.
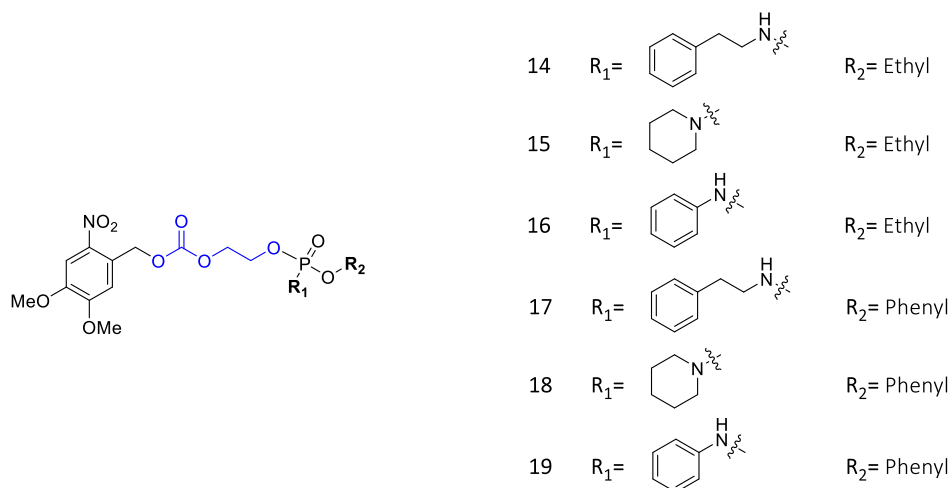
Figure 29: Model compounds **14–19** bearing amine-containing cargo.

The linkers **14**, **15** and **16** (δ_P = 10.36, 9.39, and 3.29 ppm, respectively) successfully provided intermediates **14-I** (δ_P = 10.83 ppm), **15-I** (δ_P = 9.81 ppm), and **16-I** (δ_P = 3.62 ppm), but did not further release the amine cargos (Figure 30). The ^{13}C NMR spectra of the intermediates still possessed ^{13}C–^{31}P interactions, hinting that both cargos are still attached to the phosphorus spacers. The intermediates remained even for several days, exhibiting no cyclisation followed by the cargo release.
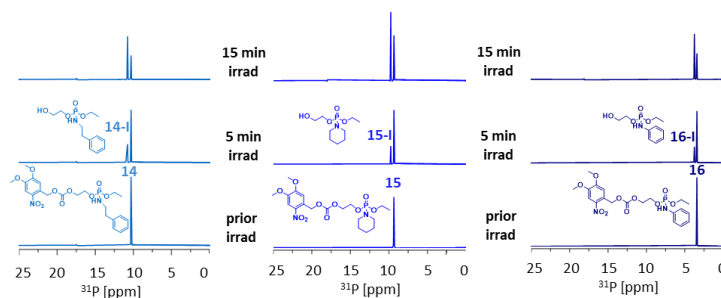
Figure 30: ^{31}P NMR spectra of linkers **14–16** (5 mM) measured before and after irradiation with UV light (365 nm) at room temperature in 50% CACO/DMSO-d_6 buffer (1:1, *v/v*; pH = 7.4).

In turn, the linkers **17** (δ_P = 5.60 ppm) and **19** (δ_P = −1.39 ppm) yielded cyclic intermediates **17-cyc-I** (δ_P = 28.05 ppm) and **19-cyc-I** (δ_P = 21.66 ppm), respectively, indicating the phenol release within 5 min of irradiation (Figure 31). Interestingly, compound **18** (δ_P = 4.55 ppm), bearing a secondary amine, showed only a trace of **18-cyc-I** (δ_P = 26.51 ppm) overnight, releasing phenol in several days. The phenol release was also detected by a missing ^{13}C–^{31}P interaction of the phenoxy group. Only aniline as a representative of the aromatic amine was released overnight, as indicated by **19-cyc-P** (δ_P = 17.56 ppm) with a slight upfield shift in the ^{31}P NMR spectrum and a disappearance of the ^{13}C–^{31}P *J*-couplings
of the amine moiety. Phenol was released much faster, already in 15 minutes (forming **19-cyc-I**). Therefore, the linkers **14–19** were inefficient in an amine-cargo release. To support the amine release, we varied the SI spacer and selected lactate.
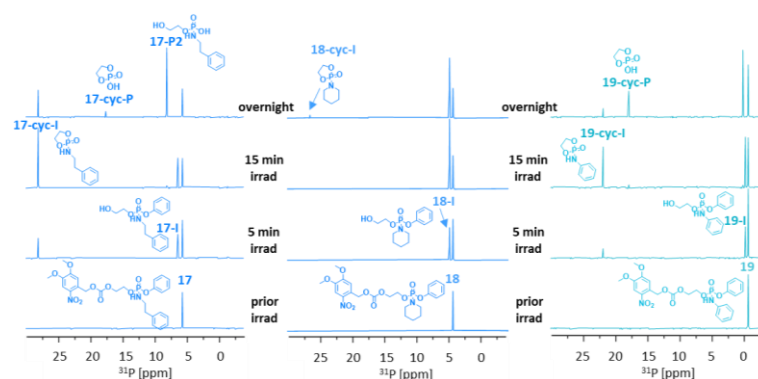
Figure 31: ^{31}P NMR spectra of linkers **17**–**19** (5 mM) measured before and after irradiation with UV light (365 nm) at room temperature in 50% CACO/DMSO-d_6 buffer (1:1, v/v; pH = 7.4).

3.2.2 Lactate Phosphate-Based Linkers

The lactate linkers **20**–**22**, structurally analogous to **14**–**16**, were prepared. The chiral lactate spacer attached to the core with a stereogenic centre on phosphorus caused the formation of two diastereoisomers detected as two ^{31}P NMR signals (Figure 32). The ^{31}P NMR spectroscopy detected successful amine release from **20** (δ_P = 9.44 and 9.84 ppm) and **21** (δ_P = 8.37 and 8.87 ppm), providing the final product **P** (δ_P = −1.05 ppm) in 5 minutes of irradiation signalled by a significant upfield ^{31}P NMR shift. However, linkers **20** and **22** partially decomposed after dissolution and yielded an unknown product (**20-X** and **22-X** with δ_P = −1.76 and −1.56 ppm, respectively) without irradiation. Both compounds have a free NH group attached to the phosphorus atom; therefore, a pH change may suppress this undesired decomposition.
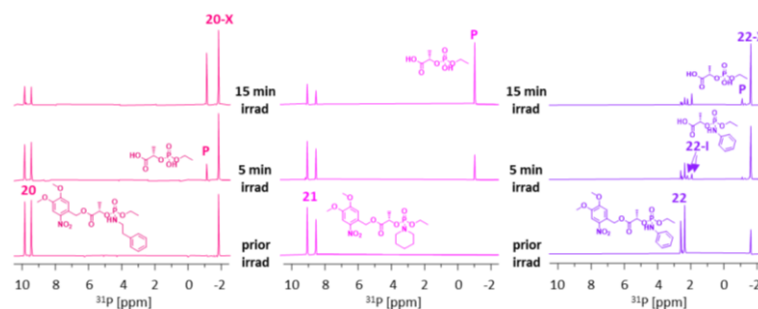
Figure 32: ^{31}P NMR spectra of linkers **20**–**22** (5 mM), measured before and after irradiation by UV light (365 nm) in a solvent mixture of 50% CACO/DMSO-d_6 (1:1, v/v; pH = 7.4) at room temperature (25 °C).

To test the pH hypothesis for eliminating the formation of the undesired products **20-X** and **22-X** in the CACO/DMSO-d_6 mixture, we screened the reactivity of **20** in several buffer mixtures. These tests showed that the undesired product **20-X** was suppressed by (i) the decrease of pH of the cacodylate buffer (to pH = 5), (ii) the alteration of the buffer system, or (iii) the change to an unbuffered environment. To retain the physiological environment, we selected 50% 2-[4-(2-hydroxyethyl)piperazin-1-yl]ethanesulfonic acid (HEPES, pH 7.4) with DMSO-d_6 system (1:1, v/v) for further investigations of linkers **20**–**22** (Figure 33).

In 50% HEPES/DMSO-d_6 solvent system, no undesired product **X** was detected prior to irradiation. Within 5 minutes of irradiation, the final product **P** (δ_P = −1.10 ppm) was detected in the cases of **20** and **21**. Compounds **20** and **21** cyclized fast; thus, we did not observe any traces of **20-I** and **21-I**, making photoactivation the rate-limiting step. For **22**, we obtained only the intermediate **22-I** (δ_P = 1.95 and 2.22 ppm) after 5 minutes of irradiation and a final product **P** was formed after 15 minutes of irradiation; we observed only traces of the undesired product **22-X** (δ_P = −1.57 ppm). In this case, NMR spectroscopy proved essential since analysis by traditionally used UV/vis spectroscopy could completely omit the formation of such undesired compound.

Figure 33: ^{31}P NMR spectra of linkers **20–22** (5 mM solutions in a solvent mixture of 50% HEPES/DMSO-d_6 (1:1, *v/v*; pH = 7.4)), measured before and after irradiation by UV light (365 nm) at room temperature (25 °C). The formation of the undesired product **X** was efficiently suppressed.

3.2.3   Characterization of the Undesired Product **X**

To identify the alternative decomposition pathway, we characterized the undesired product **X**. The single signal of ^{31}P NMR suggested that the stereogenic centre on the phosphorus atom was cleaved, and the upfield shift of this signal implied that the amine cargo was not connected to the phosphorus atom. Moreover, the chemical shifts of the undesired products were slightly different for **20-X** and **22-X** (δ_P = −1.76 and −1.56, respectively), suggesting that **20-X** and **22-X** differ, most probably in the amine moiety. Additionally, 2D experiments of **20-X** showed the key HMBC cross-peaks describing the through-bond interaction between the phenethylamine alkyl chain and lactate carbonyl (Figure 34). HR-MS confirmed the structure of the suggested carboxamide **20-X**.
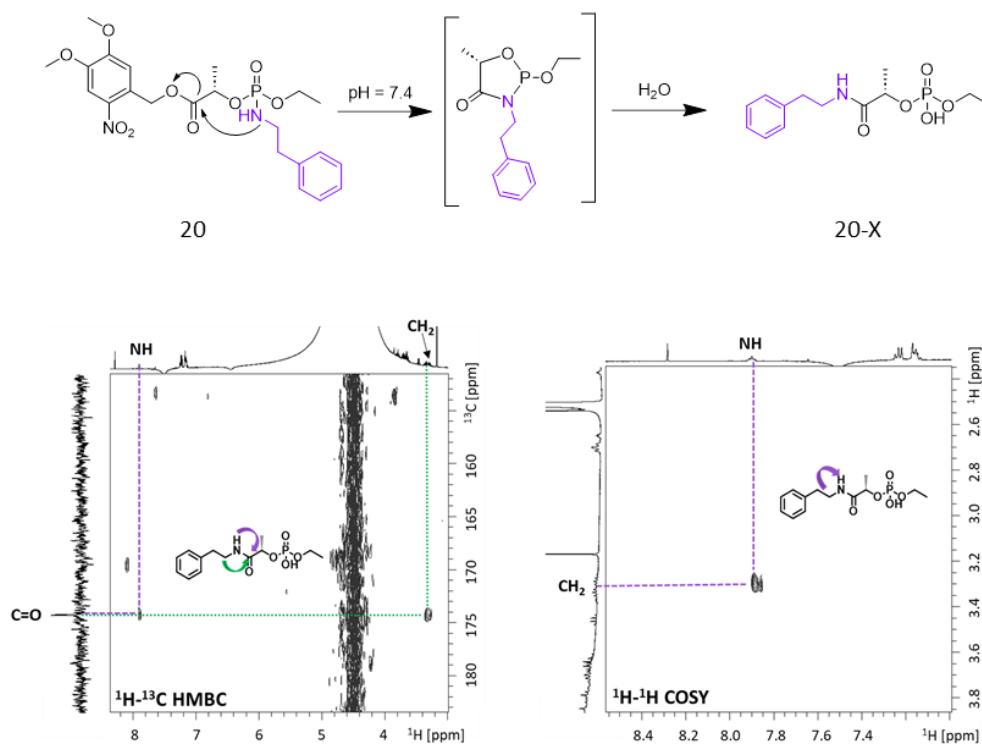
Figure 34: The proposed mechanism of the alternative decomposition of **20** to **20-X**. The structure of **20-X** was determined by ^1H–^{13}C HMBC (left) and ^1H–^1H COSY (right) spectra showing essential connections.

The proposed rearrangement was first reported by the Mulliez group.[147] To ensure that NH is crucial for the intramolecular attack of the carboxyl group, the *N*-methylated derivative of **23** was prepared (Figure 35). Upon irradiation, no undesired product **23-X** was observed in 50% CACO/DMSO-d_6 (1:1, v/v; pH = 7.4).

3.2.4 Amine Screening—Application Scope

Encouraged by the results, we prepared linkers **24–29** (Figure 35) with different cargos to further investigate the application scope of the lactate-based linkers.
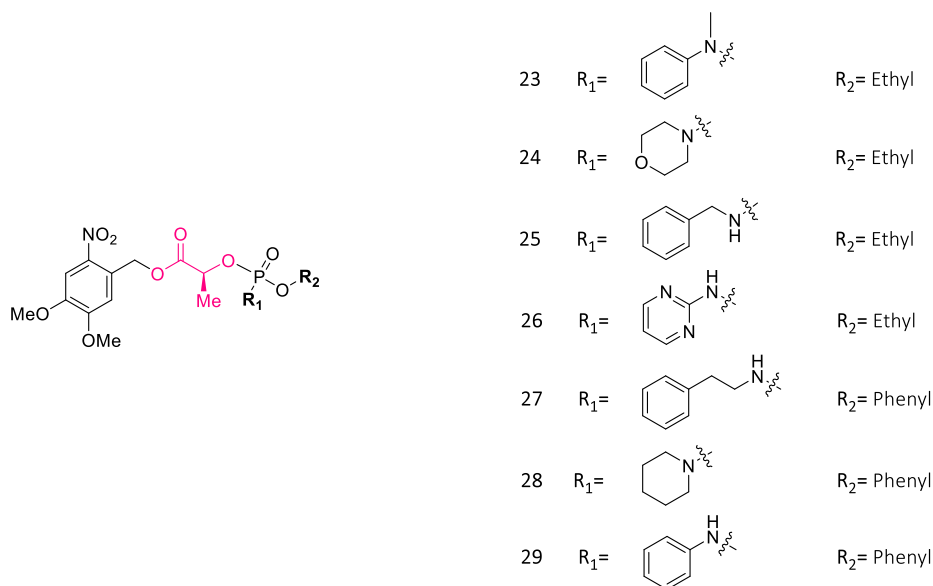
Figure 35: Lactate-based linkers **23–29**.

Upon irradiation, the linkers **23** (δ_P = 5.62 and 5.77 ppm), **24** (δ_P = 7.25 and 7.62 ppm) and **25** (δ_P = 9.25 and 9.70 ppm) provided final product **P** (δ_P = −1.04) within 5 minutes, becoming a major component within 15 minutes of irradiation. Conversely, the 2-aminopyrimidine derivative **26** (δ_P = 0.07 and 0.31 ppm) formed only an intermediate **26-I** (δ_P = −0.13 and −0.75 ppm) in 15 minutes of irradiation. The final product was formed after 19 days in the darkness (Figure 36).
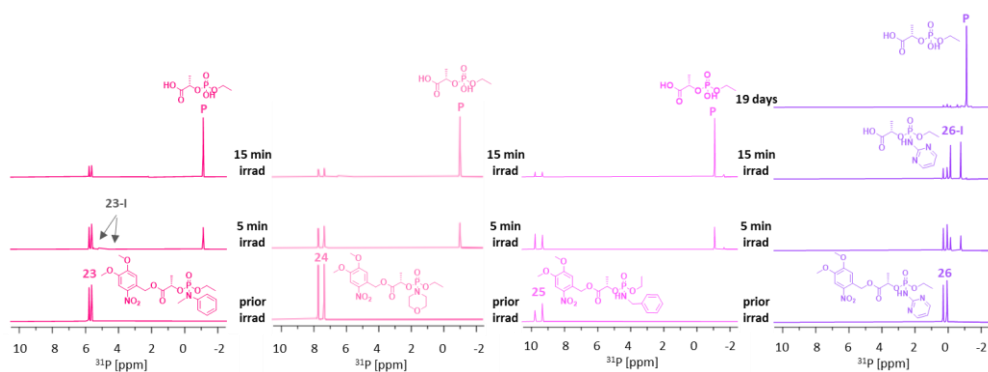
Figure 36: ^{31}P NMR spectra of linkers **23–26** (5 mM solutions in a solvent mixture of 50% HEPES/DMSO-d_6 (1:1, *v/v*; pH = 7.4)) measured before and after irradiation with UV light (365 nm) at room temperature (25 °C).

Conversely, the self-immolation of the phenyl double-cargo linkers **27** (δ_P = 4.67 and 4.94 ppm), **28** (δ_P = 3.42 and 3.79 ppm) and **29** (δ_P = −2.29 and −2.21 ppm) was slightly faster than that found in their ethyl counterparts **20–22** (Figure 37). All preferentially released the amine-cargo as recognized by detecting **27-P1**, **28-P1** and **29-P1** (δ_P =−6.24, −6.21, and −6.25 ppm, respectively). Interestingly, traces of **29-P2** (δ_P = −0.52 ppm), indicating the phenol release, were found overnight. These findings enhance the development of new drug-delivery systems offering sequential drug release for amine cargos (e.g. two drugs with synergy effect, etc.).
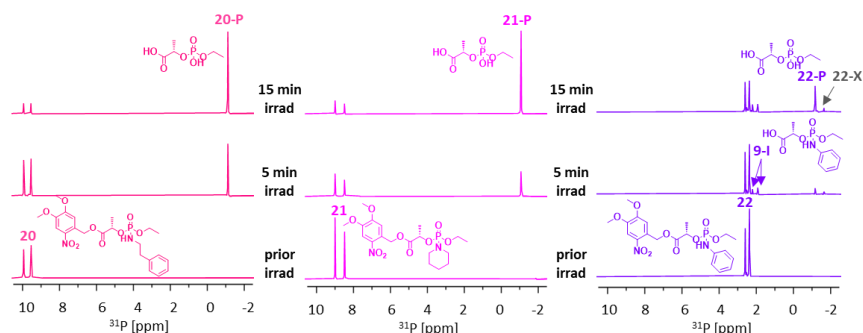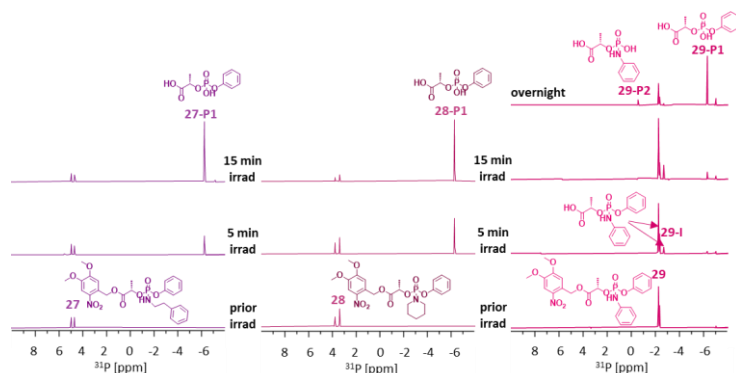

Figure 37: ^{31}P NMR spectra of linkers **27–29** (5 mM solutions in a solvent mixture of 50% HEPES/DMSO-d_6 (1:1, *v/v*; pH = 7.4)), measured before and after irradiation with UV light (365 nm) at room temperature (25 °C).

3.2.5   Conclusion

This work showed that designing a prodrug delivery system requires a detailed structure-activity relationship analysis. We successfully released amine-bearing cargos, but the study demonstrates that a universal spacer for delivering all types of amine-containing cargos will unlikely ever be designed, given the sensitivity of the phosphorus atom to substitution. ^{31}P NMR spectroscopy with *in situ* irradiation proved essential in observing all formed intermediates and products, which would otherwise be overlooked if studied by traditionally used UV/vis spectroscopy. Moreover, this setup allowed us to analyse the formed intermediates and products in real time.

^{31}P NMR parameters were crucial in the structural determination of intermediates and products of SI reactions directly in the reaction mixture. All the ^{13}C NMR signals of the moieties attached to the phosphorus were split by ^{13}C–^{31}P interaction, which helped in structural analysis qualitatively. However, the values of ^{13}C–^{31}P couplings do not follow the Karplus-like relation which aids the stereochemical analysis. Moreover, unlike ^{31}P NMR chemical shifts, ^{13}C– or ^{1}H–^{31}P *J*-couplings are studied rarely. Thus, we decided to further investigate the ^{31}P NMR parameters, which could be critical in structural analysis in a growing field of P-based compounds.

3.3 ^{31}P NMR Parameters May Facilitate the Stereochemical Analysis of Phosphorus-Containing Compounds (Paper III)

As the Introduction shows, chiral phosphorus is present in many biologically, medicinally and industrially relevant compounds. The ^{31}P NMR parameters may provide critical information that could help to solve the molecular structure and its physico-chemical properties. The incorporation of ^{13}C–^{31}P J-couplings significantly helped the structural determination of SI reaction species.

In this work, we aimed to study ^{31}P parameters (chemical shifts, RDCs, etc.). For this, we aimed to synthesise model small organic molecules with phosphorus atom incorporated into a cycle. Such compounds would be less flexible than their acyclic counterparts, which may be beneficial for searching conformational space as explained below. However, the synthesis, performed by Dr Aneta Ešnerová from the Faculty of Science, Charles University, was difficult, and several compounds were not stable enough to survive purification. Another difficulty was to separate the formed diastereoisomers from each other. Finally, we prepared and separated three pairs of diastereoisomers – **30**, **31**, and **32** (Figure 38). However, only compounds **30-SR** and **31-RR** successfully crystallized and were, thus, analysed by X-ray diffraction, which determined their absolute configuration as (*S*, *R*) and (*R*, *R*) for compound **30-SR** and **31-RR**, respectively (Figure 39). The first stereodescriptor (*S*- for **30-SR** and *R*- for **31-RR**, respectively) denotes the configuration of the proline carbon C2; the second stereodescriptor is related to the phosphorus atom. The X-ray diffraction structures were obtained by Dr Ivana Císařová from the Faculty of Science, Charles University.
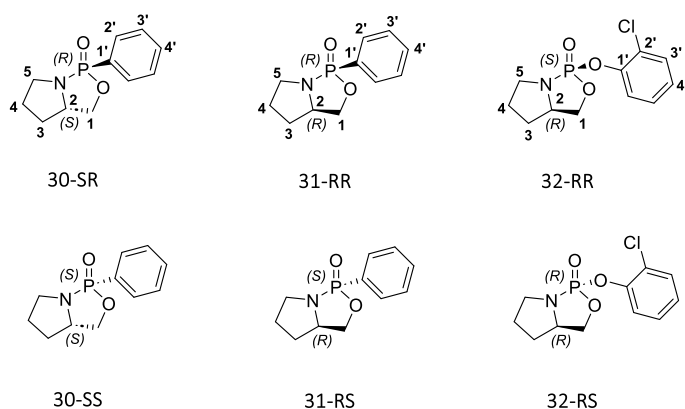
Figure 38: Model compounds **30–32** prepared for this study.

30-SR **31-RR**
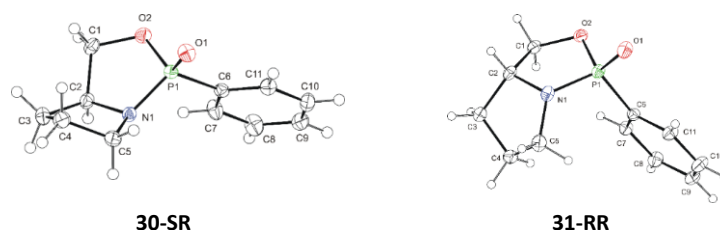
Figure 39: The X-ray structures of a) **30-SR** and b) **31-RR.**

We also designed an *ortho*-chlorophenyl phosphonate derivative of **31** due to the disturbed symmetry of the phenyl ring which should allow us to obtain more ^{1}H–^{13}C couplings to enhance the analysis. However, the synthesis was troublesome and did not provide a stable compound. Therefore, we designed and synthesized a phosphate derivative **32**. Unfortunately, compound **32** did not crystallise, thus preventing us from determining its absolute configuration. Consequently, we assigned the experimental datasets of **32** as *32-A* and *32-B* (δ_P = 21.74 and 17.01 ppm, respectively).

3.3.1 Stereochemical Investigation of **30** and **31**

First, we generated an ensemble of conformers in Schrödinger/Maestro software[148] from a single structural input. The obtained six conformers for each isomer of **30** and **31** were then optimised on a DFT level of theory with the B3LYP/DGDZVP method, and only the non-redundant structures were finally selected by analysing their dihedral angles (Figure 40).
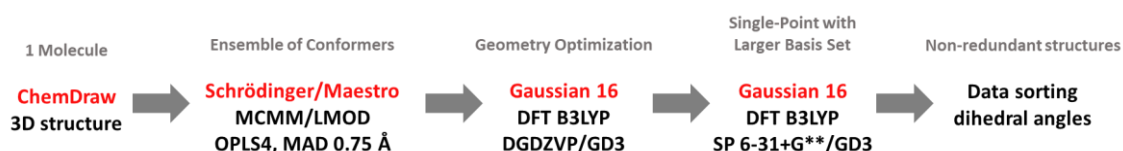
Figure 40: Schematic representation of the conformational sampling strategy used in this work.

Next, we examined the conformation of the five-membered ring using the pseudorotational cycle (Figure 41).[149, 150] The pseudorotational cycle describes all possible conformations of a ribose ring which can be considered analogous to a proline cycle of our model compounds **30** and **31**. These conformations are characterized by the phase angle *P* and maximum puckering amplitude Φ_{MAX}. The cycle is divided into twenty parts, each representing the ribose ring conformation with a corresponding *P* (0–360°).[151] The part defined by *P* = 0–36° describes the so-called North conformations, while *P* = 144–180° defines the so-called South conformations.

The configuration of prolinol governs the overall conformation of the bicyclic moiety. For **30-SR**, we found three low-energy structures: the ground-minimum one (conformer A) in the South conformation and the conformers B and C in the North conformation.

37

Furthermore, the South conformation of conformer A corresponded to the structure defined by the X-ray analysis. The Boltzmann distribution at 25 °C was established at 77, 17, and 6% for conformers A, B, and C, respectively. Figure 41 shows that the lowest-energy conformer A with phase angle $P = 17°$ possesses the $_3E$ conformation, while North conformations B ($P = 19°$) and C ($P = -35°$) are in ^3E and $_2E$ conformations, respectively.
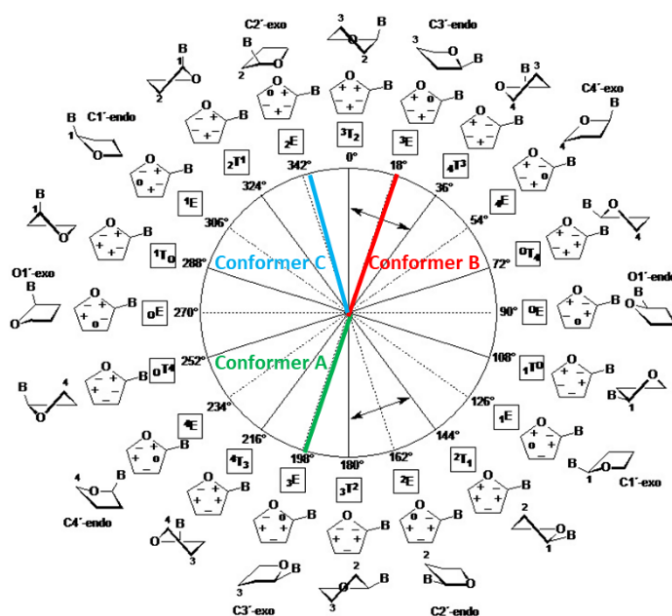
Figure 41: The pseudorotational cycle with designated conformations of **30-SR**.

We performed the same analysis for **31-RR**, finding again three conformations. The ground-minimum structure A corresponded to the X-ray analysis, and the Boltzmann distribution was defined at 47, 36, and 17% for conformers A, B, and C, respectively. The conformation A had a maximum at $P = 22°$, possessing the North conformation, while the conformers B and C ($P = -8$ and ° -37, respectively) possessed the South conformation.

We subsequently used the resulting three conformations of **30** and **31** as inputs in further analyses (*J*-couplings, RDCs).

First, we assessed whether chemical shifts might be used to determine the relative configuration of **30-SR** and **31-RR**. We calculated ^{13}C shielding constants for each of the conformations found and averaged them according to the Boltzmann distribution. Upon correlation with experimental ^{13}C chemical shifts, we obtained $R^2 > 0.99$ for both isomers in all conformations (A, B and C for both **30-SR** and **31-RR**). Hence, we could not differentiate the diastereoisomers using the ^{13}C chemical shift correlations. Based on the significant difference of ^{31}P chemical shifts of **30-SR** (δ_P = 38.67 ppm) and **31-RR** (δ_P = 34.03 ppm), we may hope to distinguish the diastereoisomers. Indeed, the calculated shielding constants were σ_P = 253.15 (**30-SR**) and 259.26 (**31-RR**), corresponding to the experimental values (a higher value of the shielding constant signifies lower shielding). However, the difference between the experimental shifts

of the two diastereoisomers is often smaller than the error of computed shielding constants, thus disabling the diastereomer discrimination. Therefore, we focused on the study of other parameters.

Therefore, we investigated the ^{13}C–^{31}P *J*-couplings connected to the stereocentre on the phosphorus atom. As shown in the Introduction, ^{31}P is highly sensitive to the chemical environment and, thus, is a promising tool for diastereoisomer discrimination. We obtained the experimental ^{13}C–^{31}P *J*-couplings from the ^{13}C APT NMR spectra measured in CDCl$_3$ and compared these experimental values with calculated *J*-couplings of **30-SR** and **31-RR**. These theoretical *J*-couplings were calculated using B3LYP/IGLO-III with empirical dispersion correction GD3 and PCM solvent model (chloroform) and then were averaged according to the Boltzmann distribution at 25 °C. Finally, the ^{13}C –^{31}P interactions provided the key information that allowed us to determine the relative configuration of **30-SR** and **31-RR** based on the interactions shown in Figure 42.
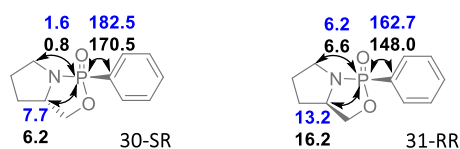
Figure 42: Experimental (blue) and calculated (black) ^{13}C-^{31}P *J*-couplings of diastereoisomers **30-SR** (left) and **31-RR** (right).

An additional tested parameter was the residual dipolar coupling in the PBLG alignment medium. We combined classically used ^{1}H–^{13}C RDCs with the ^{13}C–^{31}P ones. The experimental ^{1}H–^{13}C and ^{13}C–^{31}P RDCs and the conformers determined by the conformational analysis were used as inputs to three RDC programs (MSpin,[85] RDC@hotfcht,[87, 88] and P3D/PALES[89, 90]). To exemplify the RDC analysis, the approach is shown on **30-SR** in detail; the other compounds were treated with the same procedure.

The conformational analysis of **30-SR** presented above provided three conformers (A, B, and C). First, we entered the ^{1}H–^{13}C and ^{13}C–^{31}P experimental RDC data and the optimised conformers of **30-SR** into the MSpin software. All three conformers, A, B and C, afforded satisfactory *Q* and Pearson's correlation factors *R*: Q_{confA} = 0.0133, Q_{confB} = 0.0123, Q_{confC} = 0.0194 and R_{confA} = 0.9999, R_{confB} = 0.9998, R_{confC} = 0.9997, respectively. Then, we correlated the experimental RDC data of **30-SR** with the lowest-energy conformers A of the other three diastereoisomers (**30-SS**, **31-RS** and **31-RR**) to probe the diastereoisomer discrimination. However, the results were similar and did not allow us to determine the correct diastereoisomer using MSpin.

Unlike MSpin, RDC@hotfcht also includes experimental errors in the fitting procedure. Nevertheless, the results were like those of MSpin and did not discriminate between **30-SR** and **31-RR** diastereoisomers, nor did the software find the best-fitting conformer (Q_{RR} = 0.0202, R_{RR} = 0.9998; Q_{RS} = 0.0136, R_{RS} = 0.9999; Q_{SR} = 0.0136, R_{SR} = 0.9999; Q_{SR} = 0.0202, R_{SR} = 0.9998).

Finally, we examined the P3D software implemented in the PALES program. The P3D/PALES analysis provided promising R values of the individual conformers of **30-SR**: R_{confA} = 0.785, R_{confB} = 0.692, and R_{confC} = 0.655. Given these results, we used the lowest-energy conformer for the diastereoisomer discrimination. Unfortunately, the diastereoisomer discrimination failed to identify the correct isomer. As an example, it set the (*R*, *R*) configuration as the best-fitting structure which did not match the (*S*, *R*) configuration found by X-ray diffraction (Figure 43).
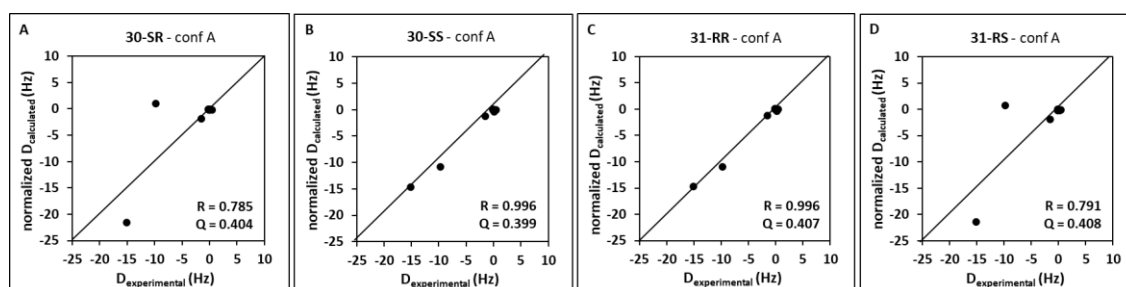
Figure 43: Correlation of experimental RDCs of **30-SR** and normalised theoretical RDC values calculated using P3D/PALES fitted to the lowest-energy conformer A of: A – **30-SR**, B – **30-SS**, C – **31-RR**, and D – **31-RS** structures.

In the RDC analysis of remaining isomers of **30** and **31**, P3D/PALES successfully discriminated **30-SS** (and **31-RR** because they are enantiomers which P3D/PALES cannot distinguish) with R ≥ 0.9, albeit failed for **31-RS** with all correlation factors above 0.8.

3.3.2 Stereochemical Investigation of **32**

The conformational sampling of **32** was complicated by its higher flexibility due to the oxygen bridge between the phosphorus atom and the C1' carbon atom of the phenyl ring. Ultimately, the conformational sampling led to more conformers than in cases of **30** and **31**. The lowest-energy conformers A of **32-RS** (population of 53%) and **32-RR** (population of 32%) were in the North conformation with a phase angle *P* of 25° and 18°, respectively. Finally, we identified nine conformers for **32-RS** and seventeen for **32-RR**. Population analysis of **32** based on the Boltzmann distribution further provided five conformers (A–E) for **32-RS** and nine (A–I) for **32-RR**. The rest of the conformers amounted to less than 2% of the total population; thus, we did not include them in weighting the NMR parameters (shielding constants and *J*-couplings).

First, we assessed whether chemical shifts might be used to determine the relative configuration of **32**. Like in the previous section, the correlations of the ^{13}C NMR chemical shifts of *32-A* (δ_P = 27.74 ppm) and *32-B* (δ_P = 17.01 ppm) and the calculated shielding constants of both diastereoisomers **32-RS** and **32-RR** did not provide decisive results as all afforded R^2 = 0.9. Unfortunately, the shielding constants of the phosphorus atom (σ_P = 274.68 for **32-RS** and 273.28 for **32-RR**) also did not discriminate the diastereomers.

In accordance with the previous analysis, we calculated the *J*-couplings for the conformers of **32**. Contrary to the results of **30-SR** and **31-RR**, the differences

in the *J*-couplings of individual isomers of **32** were negligible, not allowing for a diastereoisomer differentiation (Figure 44). Moreover, one of the key ^{13}C-^{31}P *J*–coupling between the phosphorus atom and the phenyl ring was one order of magnitude lower in **32** (ca 7 ppm) than in **30** and **31** (ca 170 ppm).
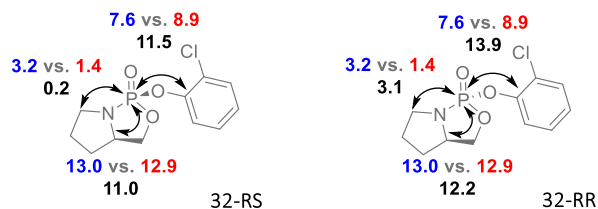
Figure 44: Experimental (blue for *32-A* dataset, red for *32-B* dataset) and calculated Boltzmann-averaged (black) ^{13}C-^{31}P *J*-couplings of the lowest energy diastereoisomers **32-RS** and **32-RR**.

The broken symmetry of the phenyl group allowed for an extraction of more ^{1}H–^{13}C and ^{13}C–^{31}P experimental RDC values from the phenyl ring than in the case of **30** and **31**. Nevertheless, no software afforded diastereoisomer discrimination of the *32-A* or *32-B* dataset, as none of the conformers provided R ≥ 0.9. The RDC analysis of flexible molecules is a prevailing problem that has yet to be solved.

3.3.3 Conclusion

We thoroughly examined the conformational space to ensure we found all possible conformations of the model P-compounds. The *J*-coupling analysis enabled us to assign the relative configuration of **30-SR** and **31-RR** using ^{13}C–^{31}P *J*-couplings. We have successfully obtained ^{1}H–^{13}C and ^{13}C–^{31}P experimental RDC values which were, together with optimized structures from conformational sampling, used as an input to three different RDC softwares. MSpin and RDC@hotfcht provided similar results for all conformations and isomers and did not allow us to determine the correct structures. P3D/PALES, designed for a diastereoisomer distinction, confirmed **31-RR** to have (*R*, *R*) configuration. However, even P3D/PALES did not provide unambiguous results in all cases. The analysis became even more difficult for flexible molecule **32**, and none of the available methods provided any insights about its configuration or conformation. We suggest that the indecisive results of RDC analysis may be caused by the small absolute sizes of the ^{13}C–^{31}P RDC values (ca 1 Hz, in comparison to ^{1}H–^{13}C RDCs which are usually in a range of −50 to 50 Hz). We also speculate that the ensemble of conformers generated via conformational sampling might not accurately represent the conformations present in the alignment medium environment since the bulky alignment medium may influence the analyte conformation. Therefore, we focused on generating a new ensemble of structures that could better describe the reality of the RDC sample.

3.4 Exploring the Impact of Alignment Media on RDC Analysis of Phosphorus-Containing Compounds: A Molecular Docking Approach (Paper IV)

In the previous chapter, we showed that the stereochemical analysis of flexible molecules is still problematic.[82] For flexible molecules, the conformational analysis results in an extensive ensemble of conformers and using the energy level difference, their population can be estimated based on the Boltzmann distribution. However, these low-energy conformers failed in the *J*-coupling and RDC analysis in the previous study. The reason may be a specific interaction of the analyte with the alignment medium, which may lead to a conformational change in the analyte molecule. Therefore, we applied a molecular docking approach to generate a more realistic ensemble of conformers in the presence of the thorough-studied PBLG alignment medium.

Molecular docking is widely used in medicinal chemistry to find the best structural match between the enzymatic active site and its potential inhibitor. It predicts the binding pose of a substrate (ligand) and the binding site of a macromolecule (a protein).[152] The binding pose designates the conformation, position, and orientation of the ligand docked. The macromolecules may change the substrate conformation with an energy penalty, which is compensated by forming the ligand-protein complex (Figure 45). In this work, we treated the small molecule as a ligand within the binding site of the PBLG alignment medium. Thus, the docked conformations may not necessarily possess ground-minimum energy.

Figure 45: A schematic representation of the molecular docking: a substrate (red) docked into a binding site of a macromolecule (green). The change of the conformation of the substrate requires energy, but the complex formation decreases the energy, making it energetically beneficial.

We applied molecular docking on three classes of compounds (Figure 46) differing in flexibility. The degree of molecular flexibility can be quantified by the $nConf_{20}$ parameter.[153] This parameter represents the number of energetically accessible conformations with energies within a selected energy range. In this work, we selected an energy threshold of 20 kcal mol^{-1} from the lowest-energy conformation. Lower values of $nConf_{20}$ designate higher rigidity of the molecule, while higher values indicate increasing flexibility. The highly rigid class of compounds (**33**, $nConf_{20}$ = 0) was represented by phosphorylated derivatives of isopinocampheol (IPC), often used as a rigid model analyte in the development of RDC analysis. [68, 70, 93, 154] Compound **33** was prepared by Mgr. Hugo Kocek (IOCB, Prague). Compounds studied in our previous work were selected as representatives for the mildly flexible (**30-SR** and **31-RR**, nConf20 = 2)

and flexible (**32-RR** and **32-RS**, nConf20 = 8) class of molecules.[155] The absolute configuration of **30**, **31** and **33** was determined by X-ray diffraction analysis in collaboration with Dr Ivana Císařová and Prof. Aleš Růžička from the Faculty of Science, Charles University (Figure 47).

nConf$_{20}$ = 0   nConf$_{20}$ = 2   nConf$_{20}$ = 8
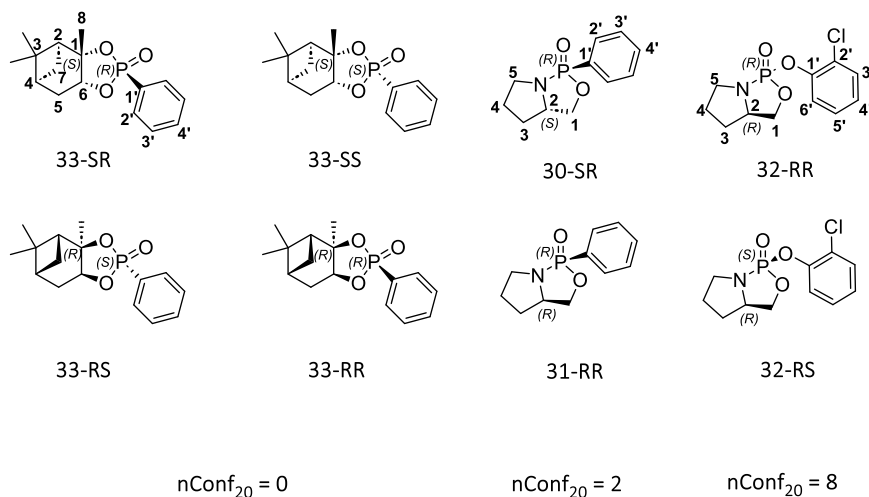
Figure 46: Chemical structures and atom numbering of the compounds studied.

a)   b)

c)   d)
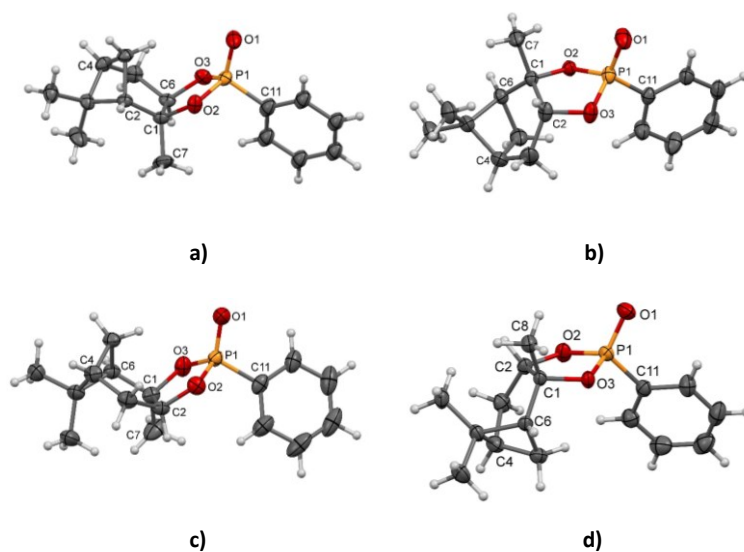
Figure 47: The X-ray structures of a) **33-SR**, b) **33-SS**, c) **33-RS**, and c) **33-RR**.

The model compounds were docked into the PBLG[92, 93, 156-159] alignment medium in Schrödinger.[160] Although the crystal structure of PBLG is not available, PBLG forms a well-defined chiral helix[91] with grooves within which the model compounds may intercalate. As a result, a new ensemble of conformers may be obtained.

Indeed, we obtained new conformer ensembles, which were then entered into P3D/PALES[89] together with experimental RDC values to retrieve theoretical back-calculated RDCs. Subsequently, we averaged the theoretical RDCs of individual conformers by the Boltzmann distribution and obtained a fit of the experimental and Boltzmann-distribution averaged RDCs qualified by Pearson's correlation factor *R*.

Additionally, we optimised the calculated docking structures by constrained optimization in internal coordinates (normal mode relaxation[161]). This method relaxes the bond lengths and molecular angles while attaining the general shape of the molecule. The workflow was then repeated for these relaxed structures. Lastly, we compared the correlation factors R from the previous low-energy (ground-minimum, GM) study, molecular docking and the normal mode relaxation (Figure 48).
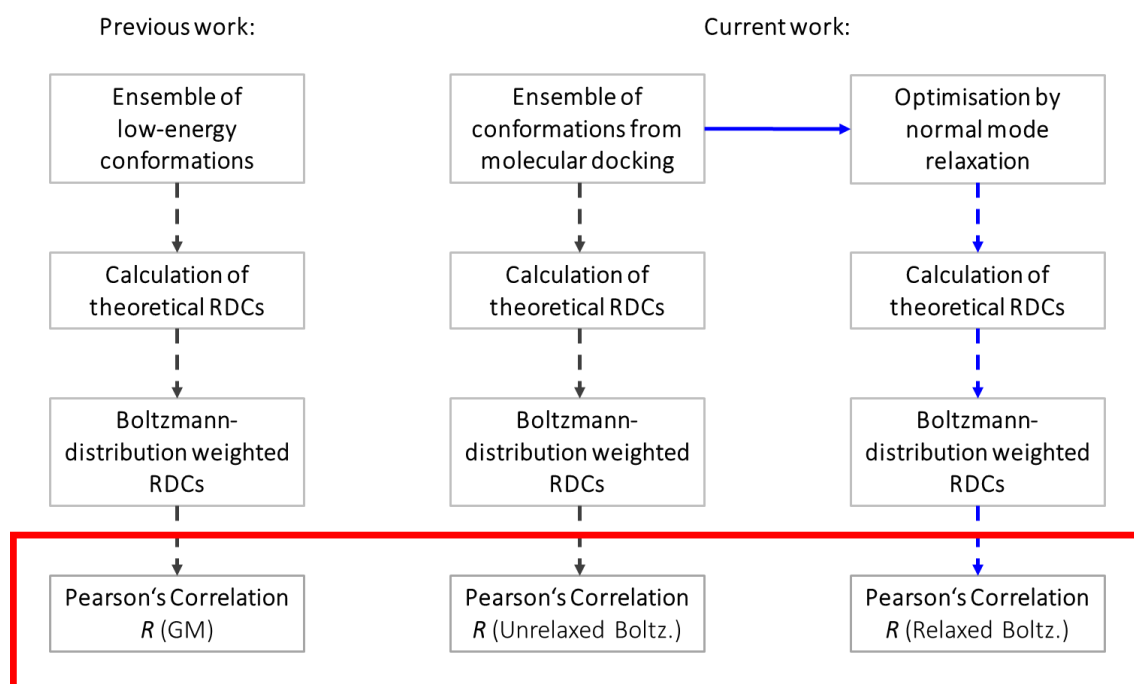
Figure 48: General workflow used in this work.

3.4.1 Molecular Docking of Rigid Molecules (**33**)

The conformational sampling, performed as shown in the previous section,[155] found only one conformation for each isomer of **33**. In turn, molecular docking revealed four conformers for **33-SR**, one for **33-SS**, three for **33-RS**, and three for **33-RR**. Interestingly, the DFT-calculated single point energies of the **33-SR** and **33-SS** conformers generated by docking were approximately 6 kcal mol^{-1} above the individual global minima. In contrast, those of **33-RS** and **33-RR** were more than 13 kcal mol^{-1} above GMs.

The reason behind the different energy levels may lie in the configuration of carbon C1 (Figure 46). The configurations **33-SR** and **33-SS** (Figures 49a and 49b) seemed to fit well into the cavity of PBLG. Interestingly, **33-SR** and **33-SS** (Figures 49c and 49d) were oriented inversely, with the P=O bond directed outside the polymer groove. In most cases, we detected a π–π stacking between the polymer and the analyte's phenyl groups with a distance of ∼ 4 Å, which is typical for π–π stacking.[162]
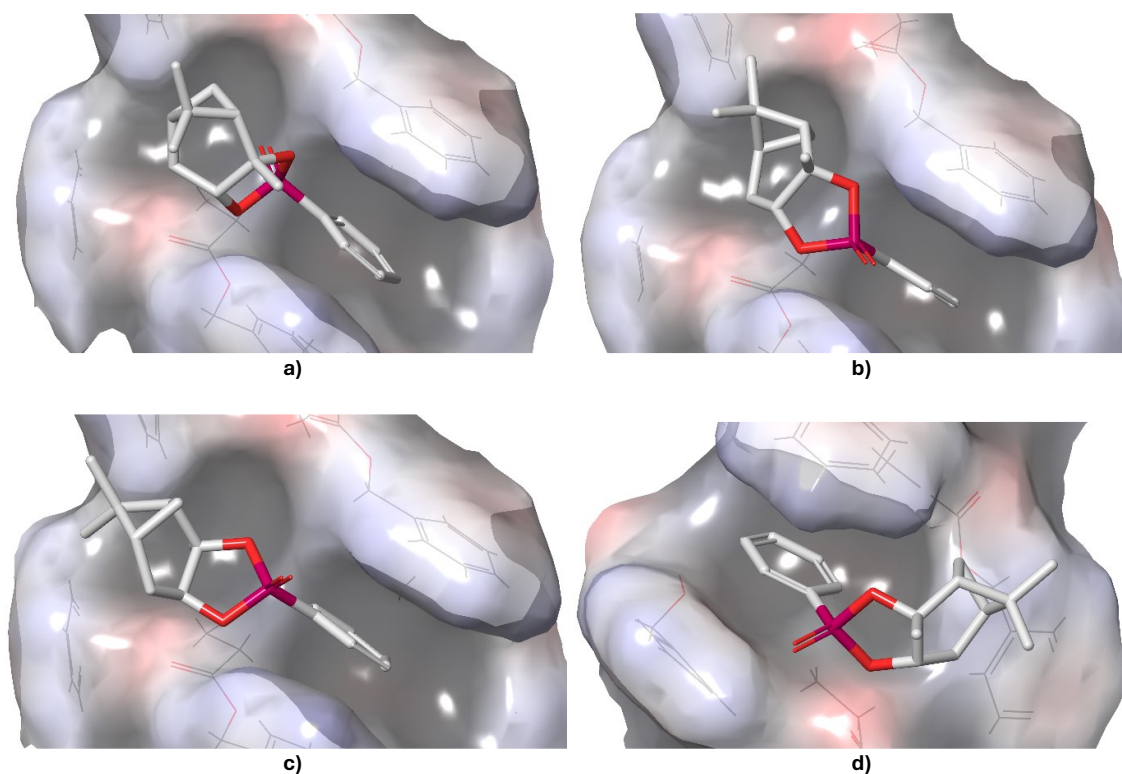
Figure 49: Docked poses of a) **33-SR**, b) **33-SS**, c) **33-RS**, and d) **33-RR** in the helical structure of PBLG obtained from molecular docking, clearly indicating different binding poses of each isomer. Carbon atoms are white, oxygen in red and phosphorus in purple; hydrogens were omitted for clarity.

Figure 50a compares the structures of **33-SR** with experimental datasets of all isomers of **33**. The correct experimental dataset *33-SR* (and its enantiomer *33-RS* as P3D/PALES cannot differentiate enantiomers) provided high *R* value (Unrelaxed Boltz.: $R_{33\text{-}SR} = 0.899$, $R_{33\text{-}RS} = 0.897$ vs $R_{33\text{-}SS} = 0.596$, $R_{33\text{-}RR} = 0.586$). Unfortunately, as shown in Figure 50b, the correlation of the structures of **33-RR** with experimental dataset of all four isomers of **33** yielded high correlation factors for the correct *33-RR* and *33-SS* datasets but even higher *R* correlation factors for the incorrect datasets *33-SR* and *33-RS* (Unrelaxed Boltz.: $R_{33\text{-}SR} = 0.947$, $R_{33\text{-}SS} = 0.842$, $R_{33\text{-}RS} = 0.961$, $R_{33\text{-}SR} = 0.799$). Therefore, we could determine the relative configuration of the diastereoisomers only unilaterally; the correct relative configuration was determined only for **33-RS/33-SR**.

To improve the diastereoisomer differentiation, we optimised the docking structures by the normal mode relaxation.[161, 163] These relaxed structures offered energies of about 1 kcal mol^{-1} for **33-SR** and **33-SS** and 2–4 kcal mol^{-1} for **33-RS** and **33-RR** above the corresponding GM, which is significantly lower than the energies of conformers from the unrelaxed structures (6 a 13 kcal mol^{-1}, respectively). Running the RDC analysis with these conformers enhanced the diastereoisomer distinction, albeit only slightly, providing no definite results (Figure 50a, Relaxed Boltz.: $R_{33\text{-}SR} = 0.911$, $R_{33\text{-}SS} = 0.596$, $R_{33\text{-}RS} = 0.909$, $R_{33\text{-}SR} = 0.578$; Figure 50b, Relaxed Boltz.: $R_{33\text{-}SR} = 0.947$, $R_{33\text{-}SS} = 0.820$, $R_{33\text{-}RS} = 0.964$, $R_{33\text{-}SR} = 0.769$).

Figure 50: Pearson correlation coefficients (R) of RDCs calculated using low-energy (GM) conformers, molecular docking (Unrelaxed Boltz.) conformers, and the same conformers after normal mode relaxation (Relaxed Boltz.) of isomers **33-SR** (a) and **33-RR** (b) correlated with *33-SR* (blue), *33-SS* (orange), *33-RS* (red), and *33-RR* (green) RDC experimental datasets.

All three approaches (low-energy, molecular docking and normal mode relaxation) gave similar results. This is probably caused by the fact that compound **33** is highly rigid and does not tend to alter its shape to fit into the grooves of the medium.

3.4.2 Molecular Docking of Mildly Rigid Molecules (**30** and **31**)

Compounds **30-SR** and **31-RR** possess higher flexibility than **33**. However, they are still relatively rigid due to the bicyclic system, as denoted by the low flexibility parameter (nConf$_{20}$=2). While the standard conformational sampling based on GM provided three low-energy conformers for each isomer of **30-SR** and **31-RR**,[155] the molecular docking approach identified tens of conformations for each isomer. We applied a machine learning-based algorithm to eliminate the redundant conformers and found 22 binding poses for **30-SR** and 21 poses for **31-RR**. The DFT single-point energies found all these conformers at 4–10 kcal mol^{-1} above the corresponding GMs. The corresponding relaxed structures were 1–6 kcal mol^{-1} above the GMs.

The experimental RDCs of *30-SR* dataset provided the best correlation with the correct **30-SR** structures in all three approaches (Figure 51a). The molecular docking and the normal mode relaxation markedly improved the correlations. Moreover, it also enhanced the diastereoisomer discrimination – GM: $R_{30\text{-SR}}$ = 0.766, $R_{31\text{-RR}}$ = 0.437; Unrelaxed Boltz.: $R_{30\text{-SR}}$ = 0.815, $R_{31\text{-RR}}$ = 0.491; Relaxed Boltz: $R_{30\text{-SR}}$ = 0.871, $R_{31\text{-RR}}$ = 0.578.

Unfortunately, the data were undecisive in the case of **31-RR** (Figure 51b). The experimental dataset *31-RR* showed high correlation factors (GM: $R_{30\text{-SR}}$ = 0.994, $R_{31\text{-RR}}$ = 0.941) with both isomers **30-SR** and **31-RR**. The docking approach did not improve the discrimination, nor did the normal mode relaxation applied – Unrelaxed Boltz.: $R_{30\text{-SR}}$ = 0.988, $R_{31\text{-RR}}$ = 0.959; Relaxed Boltz: $R_{30\text{-SR}}$ = 0.984, $R_{31\text{-RR}}$ = 0.964. Thus, the assignment

of the relative configuration of **31-RR** failed. Similarly to **33**, we reached only unilateral diastereoisomer discrimination for mildly flexible molecules **30-SR** and **31-RR**.
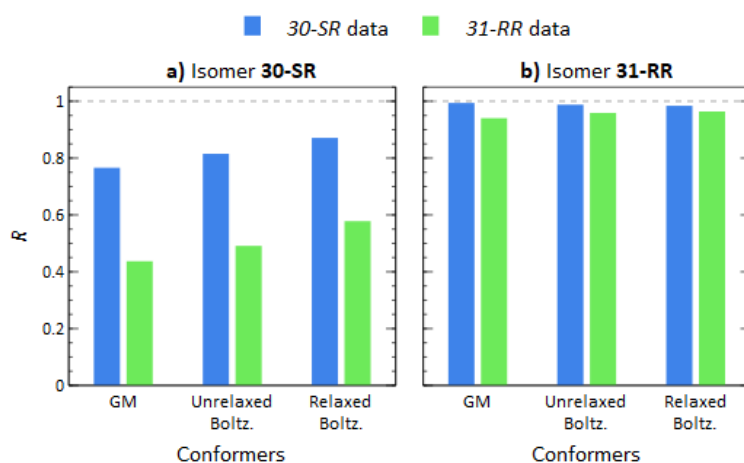
Figure 51: Pearson correlation coefficients (R) of RDCs calculated using low-energy conformers[155](GM), molecular docking (Unrelaxed Boltz.) conformers, and the same conformers after optimisation by normal mode relaxation (Relaxed Boltz.) of structures **30-SR** (a) and **31-RR** (b) correlated with the *30-SR* (blue) and *31-RR* (green) RDC experimental datasets.

3.4.3 Molecular Docking of Flexible Molecules (**32**)

Compound **32** is a flexible molecule with $nConf_{20} = 8$. The low-energy analysis identified 9 unique conformers for **32-RS** and 17 conformers for **32-RR**.[155] The molecular docking produced 23 and 34 conformations for **32-RS** and **32-RR**, respectively. These conformers were located at 6–15 kcal mol^{-1} above the individual GMs and 2–11 kcal mol^{-1} above the GMs after normal mode relaxation. Since **32** did not crystalise, we could not determine the absolute configuration.

The *32-A* experimental dataset provided higher correlation factors (R) for low-energy conformers of both **32-RS** and **32-RR** (Figure 52). For **32-RS** structures, the correlations improved when employing the docked structures and further enhanced with the relaxed conformers, ultimately reaching $R > 0.8$, declared sufficient for diastereoisomer discrimination by P3D/PALES.[89] Unfortunately, the diastereoisomer discrimination lowers with the progressing approach (Figure 52a, GM: $R_{32-A} = 0.468$, $R_{32-B} = 0.043$; Unrelaxed Boltz.: $R_{32-A} = 0.668$, $R_{32-B} = 0.428$; Relaxed Boltz: $R_{32-A} = 0.802$, $R_{32-B} = 0.590$). However, these data gave a high degree of certainty, assigning the **32-RS** structure to the *32-A* RDC experimental dataset.

For **32-RR**, the low-energy conformers provided the highest level of discrimination. However, the R remained low, $R_{32-A} = 0.582$ and $R_{32-B} = 0.117$ (Figure 52b, GM). Compared to the unrelaxed molecular-docking structures, the relaxed structures improved the correlation factors, albeit not enough for proper P3D/PALES analysis – Unrelaxed Boltz.: $R_{32-A} = 0.365$, $R_{32-B} = 0.287$; Relaxed Boltz: $R_{32-A} = 0.453$, $R_{32-B} = 0.209$.
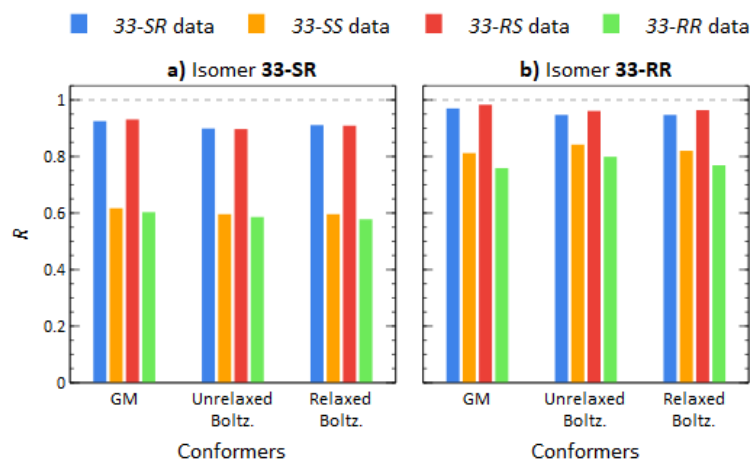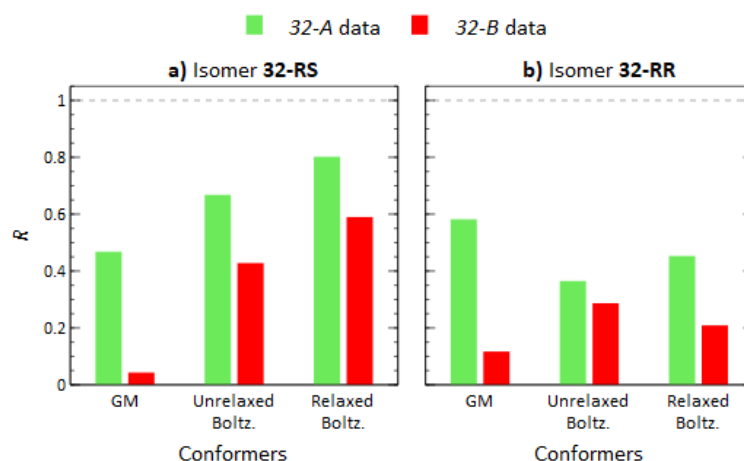
Figure 52: Pearson correlation coefficients (R) of RDCs calculated using low-energy conformers[155] (GM), molecular docking (Unrelaxed Boltz.) conformers, and the same conformers after optimisation by normal mode relaxation (Relaxed Boltz.) of structures **32-RS** (a) and **32-RR** (b) correlated with the *32-A* (green) and *32-B* (red) RDC experimental datasets.

3.4.4 Conclusion

The molecular docking results, generating new conformer ensembles presumably present in the alignment medium environment, were strongly influenced by molecular flexibility. The docking approach had little to no impact on rigid and mildly rigid molecules of **33**, **30** and **31** which may be caused by the low inclination of rigid molecules to alter their conformation. Molecular docking reflected the higher molecular flexibility of **32**, providing almost twice more conformers than the low-energy approach. However, molecular docking ultimately failed for flexible molecules **32**. Interestingly, we observed only unilateral results for the diastereoisomer discrimination in all cases. We suggest that the isomers providing high correlation factors for all RDC datasets (**33-SR**, **33-RS**, **30-SR**, **32-A**) move more freely in the presence of the alignment medium. This way, the extracted RDCs are only averaged values over many conformations and cannot give a definitive result. In turn, we suppose that the isomers not fitting any RDC dataset (**33-SS**, **33-RR**, **31-RR**, **32-B**) are affected by specific interactions with the alignment medium, which we could not involve in the molecular docking.

Possible enhancement of the diastereoisomer discrimination via the RDC analysis may lie in an advanced conformer optimisation, albeit at a higher computational-time cost. Highly valuable would be an insight into the possible specific analyte-medium interaction. However, to the best of our knowledge, this approach is unfeasible at this point because there is still no method that can visualise the alignment of the analyte within the RDC sample. We could also consider the employment of molecular dynamics. Molecular dynamics may help to describe molecular motion better, providing more accurate structures that can be further used in stereochemical NMR analysis.

3.5 Extending Molecular Dynamics with Dipolar NMR Tensors as Constraints to Chiral Phosphorus Compounds (Paper V)

Molecular dynamics (MD) can help to include molecular motion, resulting in a better description of the conformational space. Thus, we employed a recently developed MD method utilising NMR parameters - Molecular Dynamics with Orientational Constraints (MDOC)[84, 164] - which could better describe the motions of the flexible compounds. MDOC uses dipolar-coupling tensors (sometimes in combination with *J*-coupling or NOE tensors) to create trajectories of conformers. These tensorial constraints rotate the molecules studied to reach the NMR timescale in MD simulations, thus heating up the rotational degrees of freedom. Past MDOC studies were successful in stereochemical analysis of small flexible compounds,[165, 166] such as 1,4-diketone[167] ($nConf_{20} = 11$) or D-cellobiose[84] ($nConf_{20} = 61$). Still, it pointed out that even highly rigid molecules (such as strychnine, which is used as a standard in RDC studies) possess several conformations in solution.

In this work, we employed MDOC to investigate the stereochemistry of model phosphorus-based mildly flexible (**30**, **31**) and flexible molecules (**32**) from our previous works to directly compare the results from all used methods. Moreover, we complemented these molecules with a highly flexible menthol derivative **34** (with nConf20 = 78), as shown in Figure 53. Unfortunately, **34**, similar to **32**, did not crystalise. All the studied compounds were prepared by Dr Aneta Ešnerová (Faculty of Science, Charles University). The MDOC trajectories were simulated by Dr Ulrich Sternberg (Karlsruhe Institute of Technology, Germany).
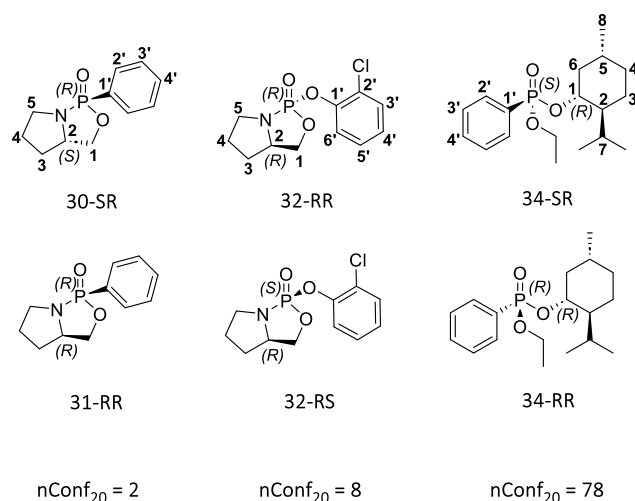
Figure 53: Chemical structure of studied compounds with their flexibility parameters.

The MDOC results were evaluated by the n/χ^2 and χ-probability.[168] The n/χ^2 is based on a known χ^2 factor, a difference between the theoretical and experimental values divided by the experimental error. To account for a different size of the datasets, the n/χ^2 factor was introduced, where *n* is the number of data points. For n/χ^2, the values greater than one designate that the theoretical values are on average within the experimental

error. On the other hand, the χ-probability factor shows how much the theoretical dataset fits the experiment on a scale from 0 to 100%, with 100% being the best fit.

3.5.1 MDOC Analysis of **30-SR** and **31-RR**

The n/χ^2 quality factors of both datasets, *30-SR* and *31-RR*, reached only 0.1 and 0.2, respectively; thus, it prevented us from determining the relative configuration (Figure 54a). This was caused by the small absolute values of $^{13}C–^{31}P$ RDCs, which did not fit their narrow error ranges. Thus, we calculated the χ-probability[168] which assigned dataset *30-SR* to **30-SR** by 90% but failed for the assignment of **31-RR** diastereoisomer (Figure 54b). Therefore, we added $^3J_{H-H}$-couplings into the MDOC simulations. The *J*-couplings were not used as additional constraints but instead calculated as the mean value of 2000 snapshots of the MDOC simulations. Complementing the experimental RDCs with theoretical $^3J_{H-H}$-couplings enhanced the n/χ^2 factors (Figure 54c) and provided the final results, assigning **30-SR** by 96% to dataset *30-SR* and **31-RR** by 97% to dataset *31-RR* (Figure 54d).
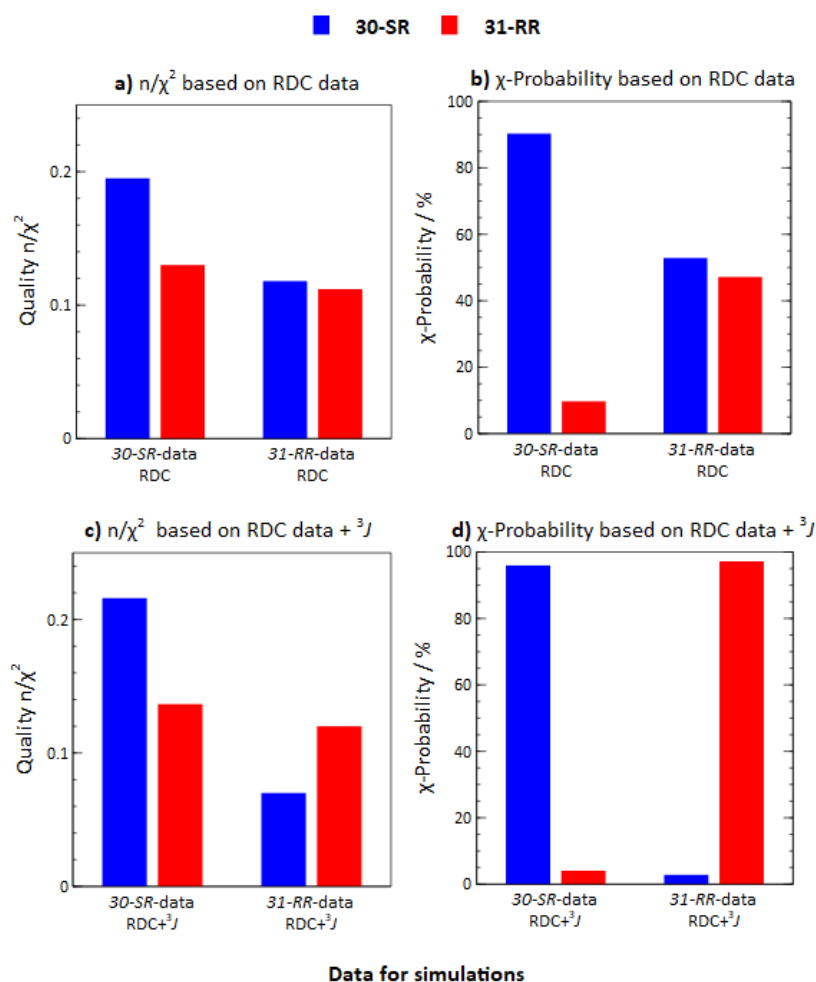
Figure 54: The MDOC simulation results of **30-SR** and **31-RR**.

MDOC simulations can also be used to search conformational space.[167] We analysed the conformational equilibria of **30-SR** using the torsion angles from the MDOC trajectories and studied the five-membered rings using the pseudorotational cycle and phase angle P.[151] Most conformations (98%) had a maximum at $P = 208°$, belonging to the South conformation. The rest (only 2%) belonged to the Noth conformation with $P = 68°$. The torsion angle about 02-C1-C2-N was by 94% in +gauche conformation with $P = 28°$ (Figure 55). In comparison to the low-energy study results of **30-SR** (Paper III, Section 3.3), the lowest-energy structure (with 77% of the population) had a maximum at $P = 17°$, belonging to the South conformation, and minor conformations B ($P = 19°$) and C ($P = −35°$) belonged to the North conformation.
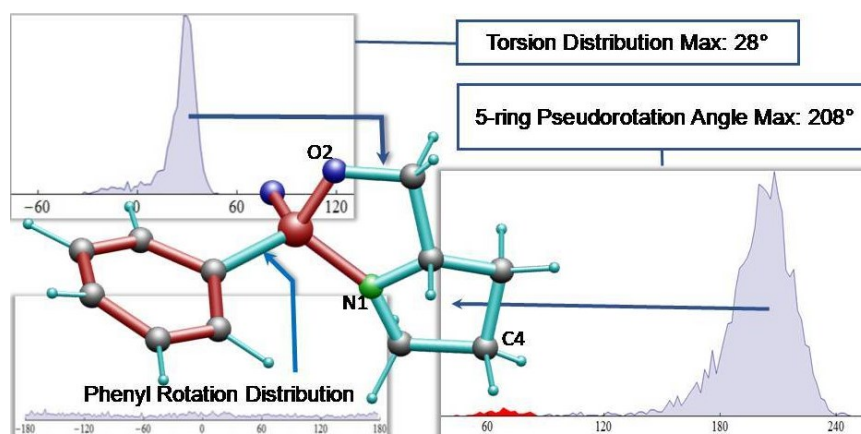
Figure 55: The conformational analysis of **30-SR**.

The conformational analysis of **31-RR** provided similar results to those of **30-SR**. Again, we identified two conformations of the five-membered ring, where the major component (98%) had $P = 168°$. The torsion angle about 02-C1-C2-N1 was by 96% in +gauche conformation with $P = 20°$. Compared to results in section 3.3 (Paper III), the Boltzmann distribution afforded three populations of 47, 36 and 17% for conformers A, B, and C, respectively. Conformer A possessed the North conformation with $P = 22°$, while conformers B and C possessed the South conformation ($P = −8°$ and $−37°$, respectively).

3.5.2 MDOC Analysis of **32-RS** and **32-RR**

Similarly to **30-SR** and **31-RR**, the quality factors n/χ^2 of the RDC datasets were low (Figure 56a), and the χ-probability based on RDC data provided only a partial result, assigning the *32-B* dataset to **32-RS** by 80% probability (Figure 56b). Thus, we calculated $^3J_{H–H}$-couplings using snapshots from the MDOC trajectories, improving the n/χ^2 (Figure 56c). The χ-probability based on RDC and theoretical *J*-couplings assigned the dataset *32-A* to **32-RR** and *32-B* to **32-RS** by 72 and 85% probability, respectively (Figure 56d). These results significantly improved over the molecular-docking study which assigned the datasets only unilaterally (*32-A* belonging to **32-RS**).
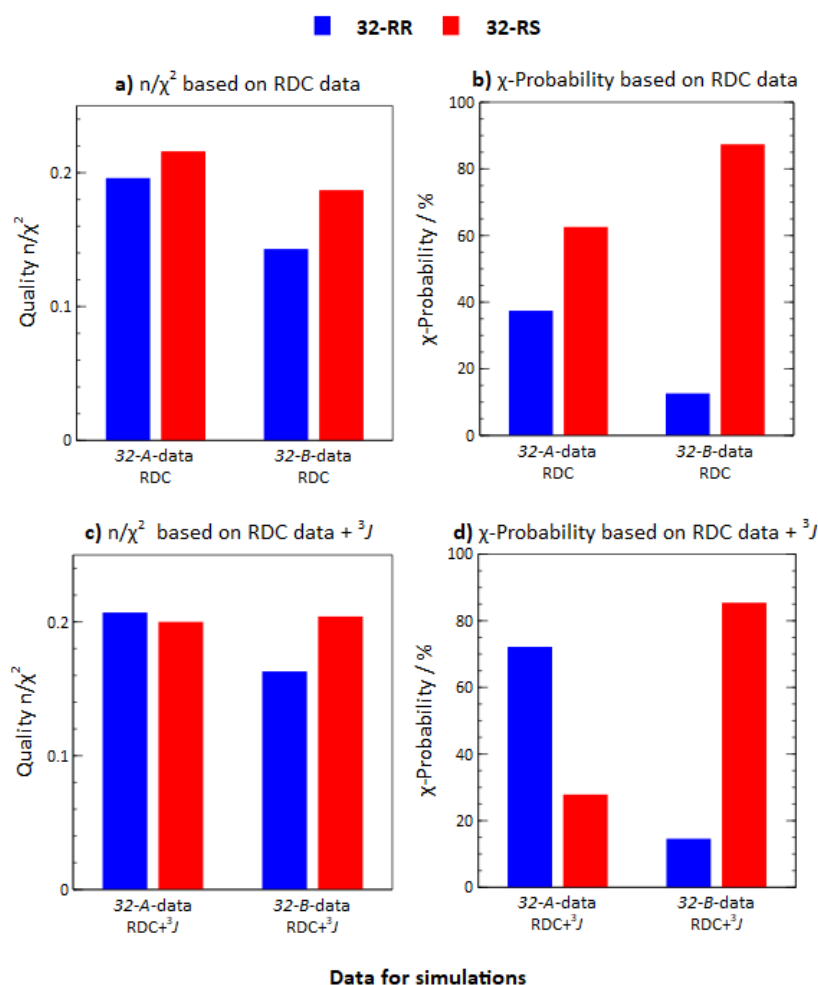
Figure 56: The MDOC simulation results of **32-RS** and **32-RR**.

The compound **32-RR** showed two conformations of the five-membered ring. The conformation with $P = 168°$ amounted to 95% of the conformational equilibrium, while the conformations with $P = 36°$ accounted only for 5%. The angle about O2–C1–C2–N1 was by 90% in +gauche conformation with $P = 22°$. The compounds **32** displayed high dynamics around the P-O1-C6 bridge. The rotation of the P-O1 bond was free, but the O1–C6 bond displayed two maxima: +100° and −100 ° (Figure 57). In contrast, the low-energy study found the lowest-energy conformer A with a population of 32% and with the phase angle $P = 18°$ (North conformation).

32-RS offered major conformation (92%) with $P = 50°$ and a minor component (8%) with $P = 178°$. The dihedral distributions matched those of **32-RR**, and the O2–C1–C2–N1 torsion angle (by 96% in +gauche) provided a maximum at $P = 30°$. Interestingly, the low-energy study offered the lowest-energy conformer A also in the North conformation with $P = 25°$ and a population of 53%.
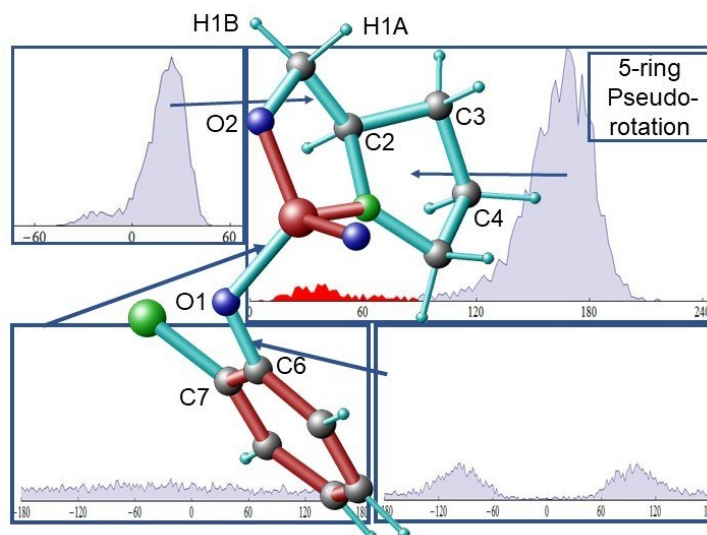
Figure 57: The conformational analysis of **32-RR**.

3.5.3   MDOC Analysis of **34-SR** and **34-RR**

The MDOC analysis of **34-SR** and **34-RR** was conducted with three experimental datasets, *34-A*, *34-B* and *34-C*. The *34-C* dataset has the $^3J_{H-H}$-couplings of the CH$_3$ group interchanged considering the *34-B* dataset because NMR cannot distinguish the protons in the CH$_3$ group. The quality factor n/χ^2 of the RDC data provided a high score; thus, we did not have to add theoretical $^3J_{H-H}$-couplings like in the previous cases (Figure 58a). The only outlier was yielded for the P-C1 RDC. This RDC value is crucial for the relative configuration assignment as it connects the chiral centre on the phosphorus atom and the C1 chiral centre on the menthol moiety. The χ-probability (Figure 58b) arrived at an 80% probability of assigning the dataset *34-A* to **34-SR**. Unfortunately, the other datasets were inconclusive, primarily because of the P-C1 outlier. The P-O1-C1 bridge is highly flexible, and the rotation around the bond is almost unrestricted. This results in an averaging of the P-C1 RDC.[169] This issue may be solved by additional constraints defining the P and C1 chiral centres, such as NOE. This is in accordance with MDOC analysis of such highly flexible compounds, e.g., D-cellobiose[84] (nConf$_{20}$ = 61) or mandelalide A[168] (nConf$_{20}$ = 947), which proved the necessity of additional NMR parameters, especially NOEs, describing the chiral centres.
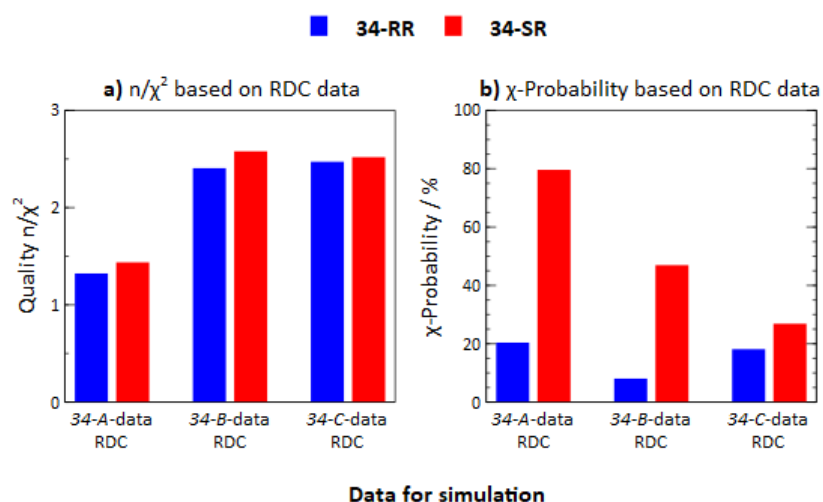
Figure 58: The MDOC simulations results of **34-RR** and **34-SR**.

3.5.4 Conclusions

MDOC approach significantly improved the results compared to the molecular docking-based approach. The MDOC simulations unambiguously determined the relative configuration of compounds **30**, **31** and **32**. Molecular docking assigned **30-SR** but failed for **31-RR**. In contrast, MDOC successfully determined the relative configuration of both structures. In the case of **32**, molecular docking did not reach satisfactory R (always $R <$ 0.8), but MDOC doubtlessly discriminated both diastereoisomers of **32**. However, the analysis failed for the highly flexible compound **34**, probably due to the rapid motion of the P-O1-C1 bridge and the P-C1 RDC outlier. Nonetheless, the χ-probability assigned the **34-SR** structure to the *34-A* dataset, reaching at least unilateral discrimination. The MDOC proved to be a helpful tool in assigning the relative configuration and determining the conformational equilibria. Nevertheless, the molecular flexibility was and still is a problem that needs further investigation.

4. Summary

In this work, I investigated the structures and physico-chemical properties of phosphorus-containing organic molecules using advanced NMR spectroscopy and theoretical methods. The studied compounds were prepared with two main aims: 1) as model compounds for structure-activity relationship NMR study of self-immolation, leading to the design of novel phosphate-based linkers for the development of new drug delivery systems (papers I and II), and 2) as model compounds for ^{31}P NMR studies where we implemented ^{31}P NMR parameters to the structural assignments using advanced NMR experiments (papers III–V).

This work investigated the self-immolation of model phosphate-based linkers with two potentially releasable cargos by UV light in real time. The ^{31}P NMR reaction monitoring combined with ^{13}C NMR data provided key information on newly formed (even metastable) reaction species, such as cyclic intermediates, proving that the cargo was released by intramolecular cyclization. We found that the pK_a of the attached cargo directs its release – the more acidic cargo is released preferentially. Following the Thorpe-Ingold effect, the bulkier SI spacer (α-hydroxybutyrate) released the cargo faster. Encouraged by this knowledge, we designed new linkers bearing two cargos and successfully released both sequentially within a tunable timeframe. Applying these results, we designed new SI linkers for the release of overall problematic amine-containing drugs. We also revealed an alternative decomposition pathway of several lactate linkers with a P-NH-R structural motif. The intramolecular rearrangement, where the NH-R moiety moved from the phosphorus atom to the lactate part, would be easily misinterpreted using the standard UV/vis method.

The ^{31}P NMR studies showed that the ^{31}P parameters can contribute to the stereochemical analysis but not as fundamentally as we expected. The ^{31}P chemical shifts and ^{13}C–^{31}P J-coupling analysis of low-energy structures successfully determined the relative configuration of mildly rigid compounds (**30-SR** and **31-RR**). On the contrary, the ^{13}C–^{31}P RDCs using the low-energy structures of **30-SR** and **31-RR** did not determine the relative configuration. Moreover, the J-coupling and RDC analysis failed in the study of flexible molecule **32** with nConf$_{20}$ = 8. Molecular flexibility became the central issue of this work. Flexible molecules may adopt different conformations due to the interaction with the alignment medium, which we simulated by molecular docking. Using molecular docking, we generated an ensemble of structures with energies in the 4–15 kcal mol^{-1} range above the respective GMs. This method improved the results over the low-energy approach, albeit mostly only unilaterally (determining the relative configuration of rigid and mildly flexible molecules **33** and **30**). Again, it failed for more flexible molecule **32**. Thus, the molecular dynamics with orientational constraints method was employed and enabled us to determine the relative configuration of mildly flexible (**30** and **31**) and flexible (**32**) molecules with a high degree of probability, in some cases reaching

almost 100%. However, the compounds with the most degrees of freedom (**34** with $nConf_{20}$ = 78) remain challenging for the stereochemical analysis and call for constant development of new NMR methods.

5. Author Contributions

Paper I

P. Šimon,[+] **M. Tichotová**,[+] M. García Gallardo, E. Procházková* and O. Baszczyňski,* *Phosphate-Based Self-Immolative Linkers for Tuneable Double Cargo Release*, Chem. Eur. J., *2021*, **27**, 12763–12775

[+]these authors contributed equally (synthesis vs NMR)

In this work, I was responsible for all the NMR experiments performed to determine the structure of newly prepared phosphate-based SI linkers and the NMR experiments with *in situ* irradiation for reaction monitoring of photoactivated self-immolation. I interpreted the NMR data, and my results directed the further design of new SI linkers with desired properties. I significantly contributed to writing the publication. (M.C. contribution 25%)

Paper II

M. Đud,[+] **M. Tichotová**,[+] E. Procházková* and O. Baszczyňski,* *Phosphate-Based Self-Immolative Linkers for the Delivery of Amine-Containing Drugs,* Molecules, *2021*, **26**, 5160

[+]these authors contributed equally (synthesis vs NMR)

Similar to the previous work, my task was recording the NMR spectra and interpreting the NMR data to investigate the structure of novel SI linkers and monitor their fragmentation activated by light. The key moment was revealing the alternative decomposition pathway by careful analysis of NMR data, leading to a change of solvent system used for amine release. I participated in writing the publication. (M.C. contribution 30%)

Paper III

M. Tichotová, A. Ešnerová, L. Tučková, L. Bednárová, I. Císařová, O. Baszczyňski and E. Procházková,* *^{31}P NMR Parameters May Facilitate the Stereochemical Analysis of Phosphorus-Containing Compounds*, J. Magn. Reson., *2022*, **336**, 107149

I was responsible for all the NMR experiments needed for the structural determination of model phosphorus compounds. I employed DFT calculations for geometry optimization and calculation of NMR parameters, collected the ^{31}P NMR parameters and compared them with DFT-calculated values. Next, I prepared the RDC samples and implemented advanced RDC experiments, such as F1-coupled HSQC and ^{2}H image experiment, at IOCB. I applied the new RDC software P3D/PALES, compared it with other RDC approaches and searched the conformational space using the Schrödinger/Maestro program package. I evaluated all the results and participated in writing the publication. (M.C. contribution 35%)

Paper IV

M. Christou Tichotová, L. Tučková, H. Kocek, A. Růžička, M. Straka and E. Procházková,* *Exploring the Impact of Alignment Media on RDC Analysis of Phosphorus-Containing Compounds: A Molecular Docking Approach*, Phys. Chem. Chem. Phys., *2024*, **26**, 2016–2024

I suggested the idea that the ensemble of low-energy conformers probably does not describe the real conformers in the presence of the alignment medium. I performed all the NMR experiments throughout this study. I carried out the RDC study consisting of extraction of experimental RDCs, sample optimisation and preparation, generation of a new ensemble of conformers based on molecular docking, calculation of theoretical RDCs and correlation of the experimental and calculated data. I significantly contributed to writing the manuscript. (M.C. contribution 35%)

Paper V

U. Sternberg, **M. Christou Tichotová**, L. Tučková, A. Ešnerová, J. Hanus, O. Baszczyňski, and E. Procházková,* *Extending Molecular Dynamics with Dipolar NMR Tensors as Constraints to Chiral Phosphorus Compounds*, Phys. Chem. Chem. Phys., *2024, under revision*

I initiated the collaboration with Dr Sternberg at the ENC conference in California to apply the MDOC method to our model compounds. I was responsible for all the NMR experiments and the structural determination of model phosphorus compounds. I performed the RDC experiments, optimised and prepared the sample and extracted the experimental data. I significantly contributed to writing the publication. (M.C. contribution 30%)

6. References

1   J. W. F. Nicholls, J. P. Chin, T. A. Williams, T. M. Lenton, V. O'Flaherty and J. W. McGrath, *Front. Microbiol.*, 2023, **14**, 1239189.

2   B. L. Hess, S. Piazolo and J. Harvey, *Nat. Commun.*, 2021, **12**, 1535.

3   M. P. Brady, R. Tostevin and N. J. Tosca, *Nat. Commun.*, 2022, **13**, 5162.

4   F. H. Westheimer, *Phosphorus Chemistry*, American Chemical Society, 1992, **486**, 1, 1-17.

5   S. Fukumoto, *BoneKEy Rep.*, 2014, **3**, 497.

6   A. C. Kogawa, B. G. Cernic, L. G. D. do Couto and H. R. N. Salgado, *Saudi Pharm. J.*, 2017, **25**, 934-938.

7   A. Sharpley, H. Jarvie, D. Flaten and P. Kleinman, *J. Environ. Qual.*, 2018, **47**, 774-777.

8   K. Picard, M. Griffiths, P. A. Senior, D. R. Mager and C. Richard, *J. Ren. Nutr.*, 2023, **33**, 443-449.

9   I. Sádlíková, F. Buňka, P. Budinský, V. Barbora, V. Pavlínek and I. Hoza, *LWT - Food Sci. Technol.*, 2010, **43**, 1220-1225.

10 C. Zhang, Y. Jiang, S. Li, Z. Huang, X. Q. Zhan, N. Ma and F. C. Tsai, *Heliyon*, 2022, **8**, e11225.

11 M. I. Abdullin, A. F. Valeyev, V. P. Biryukov, G. M. Trosman and K. S. Minsker, *Polym. Sci. (U.S.S.R.)*, 1985, **27**, 105-111.

12 L. G. Costa, *Toxicol. Sci.*, 2017, **162**, 24-35.

13 S. A. Hall, *Agricultural Control Chemicals*, American Chemical Society, 1950, **1**, 29, 150-159.

14 D. W. Johnson, *Advances in Tribology*, InTechOpen, 2016, 8, 175-195.

15 N. K. Batchu, Z. Li, B. Verbelen and K. Binnemans, *Sep. Purif. Technol.*, 2021, **255**, 117711.

16 D. Joly, D. Tondelier, V. Deborde, W. Delaunay, A. Thomas, K. Bhanuprakash, B. Geffroy, M. Hissler and R. Réau, *Adv. Funct. Mater.*, 2012, **22**, 567-576.

17 D. Guelfi, A. P. P. Nunes, L. F. Sarkis and D. P. Oliveira, *Sustainability*, 2022, **14**, 14266.

18 Unknown, *Gard*, 1841.

19 M. Walsh, G. Schenk and S. Schmidt, *J. Sustain. Agr.*, 2023, **1**, 2.

20 D. Cordell and S. White, *Annu. Rev. Environ. Resour.*, 2014, **39**, 161-188.

21 T. S. Elliott, A. Slowey, Y. Ye and S. J. Conway, *MedChemComm*, 2012, **3**, 735-751.

22 H. Yu, H. Yang, E. Shi and W. Tang, *Med. Drug. Discov.*, 2020, **8**, 100063.

23 M. Krečmerová, P. Majer, R. Rais and B. S. Slusher, *Front. Chem.*, 2022, **10**, 889737.

24 J. B. Rodriguez and C. Gallo-Rodriguez, *ChemMedChem*, 2019, **14**, 190-216.

25 Y. Mehellou, H. S. Rattan and J. Balzarini, *J. Med. Chem.*, 2018, **61**, 2211-2226.

26 K. A. Lyseng-Williamson, N. A. Reynolds and G. L. Plosker, *Drugs*, 2005, **65**, 413-432.

27 A. Holy, *Curr. Pharm. Des.*, 2003, **9**, 2567-2592.

28 P. Pagliano, C. Sellitto, G. Scarpati, T. Ascione, V. Conti, G. Franci, O. Piazza and A. Filippelli, *Expert Opin. Drug Discov.*, 2022, **17**, 9-18.

29 M. Ghara, H. Mondal, R. Pal and P. K. Chattaraj, *J. Phys. Chem. A*, 2023, **127**, 4561-4582.

30 T. R. Puleo, S. J. Sujansky, S. E. Wright and J. S. Bandar, *Chem. Eur. J.*, 2021, **27**, 4216-4229.

31 R. Noyori and H. Takaya, *Acc. Chem. Res.*, 1990, **23**, 345-350.

32 H. U. Blaser, B. Pugin and F. Spindler, *Helv. Chim. Acta*, 2021, **104**, e2000192.

33 H.-U. Blaser, B. Pugin, F. Spindler, E. Mejía and A. Togni, *Privileged Chiral Ligands and Catalysts*, Wiley-VCH Verlag GmbH & Co. KGaA, 2011, 3, 93-136.

34 O. Mitsunobu, *Synthesis*, 1981, **1981**, 1-28.

35 C. Xie, A. J. Smaligo, X.-R. Song and O. Kwon, *ACS Cent. Sci.*, 2021, **7**, 536-558.

36 G. Wittig and U. Schöllkopf, *Chem. Ber.*, 1954, **87**, 1318-1330.

37 L. Chen, J.-T. Ren and Z.-Y. Yuan, *Green Chem.*, 2022, **24**, 713-747.

38 S. Hajimohammadi, M. Balali-Mood, L. Etemad and M. Moshiri, *Encyclopedia of Toxicology (Fourth Edition)*, Academic Press, 2024, **8**, 421-433.

39 M. Noga and K. Jurowski, *Arch. Toxicol.*, 2023, **97**, 651-661.

40 S. Mukherjee and R. D. Gupta, *J. Toxicol.*, 2020, **2020**, 3007984.

41 T. Imamoto, *Chem. Rec.*, 2016, **16**, 2659-2673.

42 A. Klaebe, A. Carrelhas, J.-F. Brazier, D. Houalla and R. Wolf, *Phosphorus Sulfur Relat. Elem.*, 1977, **3**, 61-76.

43 T. Imamoto, *Proc. Jpn. Acad. Ser. B Phys. Biol. Sci.*, 2021, **97**, 520-542.

44 W. Rehman, L. M. Arfons and H. M. Lazarus, *Ther. Adv. Hematol.*, 2011, **2**, 291-308.

45 V. Dragojlovic, *ChemTexts*, 2015, **1**, 14.

46 J. R. Deschamps, *Life Sci.*, 2010, **86**, 585-589.

47 L. Yu, S. M. Reutzel-Edens and C. A. Mitchell, *Org. Process Res. Dev.*, 2000, **4**, 396-402.

48 A. G. Petrovic, N. Berova and J. L. Alonso-Gómez, *Structure Elucidation in Organic Chemistry*, Wiley-VCH Verlag GmbH & Co. KGaA, 2015, 3, 65-104.

49 F. M. d. S. Junior and J. M. B. Junior, *Chiral Separations and Stereochemical Elucidation*, Wiley-VCH Verlag GmbH & Co. KGaA, 2023, 14, 551-591.

50 F. J. Devlin, P. J. Stephens, J. R. Cheeseman and M. J. Frisch, *J. Am. Chem. Soc.*, 1996, **118**, 6327-6328.

51 J. Mattiat and S. Luber, *Chem. Phys.*, 2019, **527**, 110464.

52 T. Fujisawa and M. Unno, *Molecular and Laser Spectroscopy*, Elsevier, 2020, **2**, 2, 41-82.

53 H. Günther, *NMR Spectroscopy: Basic Principles, Concepts and Applications in Chemistry*, Wiley-VCH Verlag GmbH & Co. KGaA, 2013.

54 P. J. Hore, *Nuclear Magnetic Resonance*, Oxford University Press, 2015.

55 M. Karplus, *J. Am. Chem. Soc.*, 1963, **85**, 2870-2871.

56 A. W. Overhauser, *Phys. Rev.*, 1953, **92**, 411-415.

57 M. Tichotová, T. Landovský, J. Lang, S. Jeziorowski, V. Schmidts, M. Kohout, M. Babor, P. Lhoták, C. M. Thiele and H. Dvořáková, *J. Org. Chem.*, 2023, DOI: 10.1021/acs.joc.2c02594.

58 A. Friebel, E. von Harbou, K. Münnemann and H. Hasse, *Ind. Eng. Chem. Res.*, 2019, **58**, 18125-18133.

59 T. Maschmeyer, L. P. E. Yunker and J. E. Hein, *React. Chem. Eng.*, 2022, **7**, 1061-1072.

60   C. Feldmeier, H. Bartling, K. Magerl and R. M. Gschwind, *Angew. Chem. Int. Ed.*, 2015, **54**, 1347-1351.

61   A. Hruzíková, L. Mužíková Čechová, S. Štěpánová, L. Tučková, M. Tichotová, A. Růžička, V. Kašička and E. Procházková, *Dyes Pigm.*, 2023, **212**, 111099.

62   E. Procházková, J. Filo, L. M. Čechová, M. Dračínský, I. Císařová, Z. Janeba, I. Kawamura, A. Naito, I. Kuběna, P. Nádaždy, P. Šiffalovič and M. Cigáň, *Dyes Pigm.*, 2022, **199**, 110066.

63   L. Čechová, J. Filo, M. Dračínský, C. Slavov, D. Sun, Z. Janeba, T. Slanina, J. Wachtveitl, E. Procházková and M. Cigáň, *Angew. Chem. Int. Ed.*, 2020, **59**, 15590-15594.

64   D. Parker, *Chem. Rev.*, 1991, **91**, 1441-1457.

65   J. M. Seco, E. Quiñoá and R. Riguera, *Chem. Rev.*, 2004, **104**, 17-118.

66   P. L. Rinaldi, *Prog. Nucl. Magn. Reson. Spectrosc.*, 1982, **15**, 291-352.

67   A. P. Kumar, D. Jin and Y.-I. Lee, *Appl. Spectrosc. Rev.*, 2009, **44**, 267-316.

68   S. Jeziorowski and C. Thiele, *Chem. Eur. J.*, 2018, **24**, 15631- 15637.

69   S. Hansmann, V. Schmidts and C. M. Thiele, *Chem. Eur. J.*, 2017, **23**, 9114-9121.

70   S. Hansmann, T. Larem and C. M. Thiele, *Eur. J. Org. Chem.*, 2016, **2016**, 1324-1329.

71   Y. E. Moskalenko, V. Bagutski and C. M. Thiele, *Chem. Commun.*, 2017, **53**, 95-98.

72   M. E. Garcia, S. R. Woodruff, E. Hellemann, N. V. Tsarevsky and R. R. Gil, *Magn. Reson. Chem.*, 2016, **55**, 206-209.

73   G. Kummerlöwe and B. Luy, *TrAC, Trends Anal. Chem.*, 2009, **28**, 483-493.

74   I. Canet, A. Meddour, J. Courtieu, J. L. Canet and J. Salaun, *J. Am. Chem. Soc.*, 1994, **116**, 2155-2156.

75   P. Trigo-Mouriño, C. Merle, M. R. M. Koos, B. Luy and R. R. Gil, *Chem. Eur. J.*, 2013, **19**, 7013-7019.

76   D. J. Philp, C. Naumann and P. W. Kuchel, *Concepts Magn. Reson. A*, 2012, **40A**, 90-99.

77   C. M. Thiele, *Concepts Magn. Reson. A*, 2007, **30A**, 65-80.

78   I. Timári, L. Kaltschnee, M. H. Raics, F. Roth, N. G. A. Bell, R. W. Adams, M. Nilsson, D. Uhrín, G. A. Morris, C. M. Thiele and K. E. Kövér, *RSC Adv.*, 2016, **6**, 87848-87855.

79   A. Enthart, J. C. Freudenberger, J. Furrer, H. Kessler and B. Luy, *J. Magn. Reson.*, 2008, **192**, 314-322.

80   N. Marcó, A. A. Souza, P. Nolis, R. R. Gil and T. Parella, *J. Magn. Reson.*, 2017, 54-60.

81   L. Yao and B. Vogeli, *Residual Dipolar Couplings: Principles and Applications*, Royal Society of Chemistry, 2024.

82   C. M. Thiele, V. Schmidts, B. Böttcher, I. Louzao, R. Berger, A. Maliniak and B. Stevensson, *Angew. Chem. Int. Ed.*, 2009, **48**, 6708-6712.

83   F. A. Roth, V. Schmidts, J. Rettig and C. M. Thiele, *Phys. Chem. Chem. Phys.*, 2022, **24**, 281-286.

84   M. Di Pietro, U. Sternberg and B. Luy, *J. Phys. Chem. B*, 2019, **123**, 8480–8491.

85   A. Navarro-Vázquez, *Magn. Reson. Chem.*, 2012, **50**, S73-S79.

86   Y. Liu, A. Navarro-Vázquez, R. R. Gil, C. Griesinger, G. E. Martin and R. T. Williamson, *Nat. Protoc.*, 2019, **14**, 217–247.

87 V. Schmidts, *Entwicklung einer Auswertungssoftware zur Anwendung Residualer Dipolarer Kopplungen in der organischen Strukturaufklärung*, Dissertation, Technische Universität Darmstadt, 2013.

88 R. Berger, C. Fischer and M. Klessinger, *J. Phys. Chem. A*, 1998, **102**, 7157-7167.

89 A. Ibáñez de Opakua, F. Klama, I. E. Ndukwe, G. E. Martin, R. T. Williamson and M. Zweckstetter, *Angew. Chem. Int. Ed.*, 2020, **59**, 6172-6176.

90 M. Zweckstetter, *Nat. Protoc.*, 2008, **3**, 679-690.

91 P. Doty, J. H. Bradbury and A. M. Holtzer, *J. Am. Chem. Soc.*, 1956, **78**, 947-954.

92 A. Marx, B. Böttcher and C. M. Thiele, *Chem. Eur. J.*, 2010, **16**, 1656-1663.

93 A. Marx and C. Thiele, *Chem. Eur. J.*, 2009, **15**, 254-260.

94 L. Chan, G. M. Morris and G. R. Hutchison, *J. Chem. Theory Comput.*, 2021, **17**, 2099-2106.

95 G. Poli, T. Seidel and T. Langer, *Front. Chem.*, 2018, **6**, 229.

96 O. Guner, O. Clement and Y. Kurogi, *Curr. Med. Chem.*, 2004, **11**, 2991-3005.

97 M. McGann, *J. Comput. Aided Mol. Des.*, 2012, **26**, 897-906.

98 H. Kim, C. Jang, D. K. Yadav and M.-h. Kim, *J. Cheminform.*, 2017, **9**, 21.

99 P. C. D. Hawkins, *J. Chem. Inf. Model.*, 2017, **57**, 1747-1756.

100 K. S. Watts, P. Dalal, R. B. Murphy, W. Sherman, R. A. Friesner and J. C. Shelley, *J. Chem. Inf. Model.*, 2010, **50**, 534-546.

101 S. Riniker and G. A. Landrum, *J. Chem. Inf. Model.*, 2015, **55**, 2562-2574.

102 S. A. Hollingsworth and R. O. Dror, *Neuron*, 2018, **99**, 1129-1143.

103 I. L. Rusakova, *Magnetochemistry*, 2022, **8**, 50.

104 D. R. Hartree and W. Hartree, *Proc. R. Soc. Lond., Ser. A Math. Phys. Sci.*, 1935, **150**, 9-33.

105 J. C. Slater, *Phys. Rev.*, 1951, **81**, 385-390.

106 T. van Mourik, M. Bühl and M. P. Gaigeot, *Philos. Trans. A, Math. Phys. Eng. Sci.*, 2014, **372**, 20120488.

107 P. Hohenberg and W. Kohn, *Phys. Rev.*, 1964, **136**, B864-B871.

108 W. Kohn and L. J. Sham, *Phys. Rev.*, 1965, **140**, A1133-A1138.

109 S. H. Vosko, L. Wilk and M. Nusair, *Can. J. Phys.*, 1980, **58**, 1200-1211.

110 A. D. Becke, *Phys. Rev. A*, 1988, **38**, 3098-3100.

111 A. D. Becke, *J. Chem. Phys.*, 1993, **98**, 5648-5652.

112 J. Tao, J. P. Perdew, V. N. Staroverov and G. E. Scuseria, *Phys. Rev. Lett.*, 2003, **91**, 146401.

113 M. M. Francl, W. J. Pietro, W. J. Hehre, J. S. Binkley, M. S. Gordon, D. J. DeFrees and J. A. Pople, *J. Chem. Phys.*, 1982, **77**, 3654-3665.

114 R. Krishnan, J. S. Binkley, R. Seeger and J. A. Pople, *J. Chem. Phys.*, 1980, **72**, 650-654.

115 W. J. Hehre, R. Ditchfield and J. A. Pople, *J. Chem. Phys.*, 1972, **56**, 2257-2261.

116 R. Ditchfield, W. J. Hehre and J. A. Pople, *J. Chem. Phys.*, 1971, **54**, 724-728.

117 M. D. Newton, W. A. Lathan, W. J. Hehre and J. A. Pople, *J. Chem. Phys.*, 1969, **51**, 3927-3932.

118 J. Fukal, O. Páv, M. Buděšínský, J. Šebera and V. Sychrovský, *Phys. Chem. Chem. Phys.*, 2017, **19**, 31830-31841.

119 S. Grimme, J. Antony, S. Ehrlich and H. Krieg, *J. Chem. Phys.*, 2010, **132**, 154104.

120 S. Miertuš, E. Scrocco and J. Tomasi, *Chem. Phys.*, 1981, **55**, 117-129.

121 A. Klamt, *J. Phys. Chem.*, 1995, **99**, 2224-2235.

122   J. Zhang, H. Zhang, T. Wu, Q. Wang and D. van der Spoel, *J. Chem. Theory Comput.*, 2017, **13**, 1034-1043.

123   K. Wolinski, J. F. Hinton and P. Pulay, *J. Am. Chem. Soc.*, 1990, **112**, 8251-8260.

124   W. Kutzelnigg, U. Fleischer and M. Schindler, *The IGLO-Method: Ab-initio Calculation and Interpretation of NMR Chemical Shifts and Magnetic Susceptibilities*, Berlin, Heidelberg, 1991.

125   A. E. A. Fouda and N. A. Besley, *Theor. Chem. Acc.*, 2017, **137**, 6.

126   M. Rezaei-Sameti, *J. Mol. Struct.-THEOCHEM*, 2008, **867**, 122-124.

127   B. Maryasin and H. Zipse, *Phys. Chem. Chem. Phys.*, 2011, **13**, 5150-5158.

128   M. Lingwood, J. R. Hammond, D. A. Hrovat, J. M. Mayer and W. T. Borden, *J. Chem. Theory Comput.*, 2006, **2**, 740-745.

129   S. K. Latypov, F. M. Polyancev, D. G. Yakhvarov and O. G. Sinyashin, *Phys. Chem. Chem. Phys.*, 2015, **17**, 6976-6987.

130   L. B. Krivdin, *Magn. Reson. Chem.*, 2020, **58**, 500-511.

131   A. Alouane, R. Labruère, T. Le Saux, F. Schmidt and L. Jullien, *Angew. Chem. Int. Ed.*, 2015, **54**, 7492-7509.

132   O. Shelef, S. Gnaim and D. Shabat, *J. Am. Chem. Soc.*, 2021, **143**, 21177-21188.

133   R. V. Gonzaga, L. A. do Nascimento, S. S. Santos, B. A. Machado Sanches, J. Giarolla and E. I. Ferreira, *J. Pharm. Sci.*, 2020, **109**, 3262-3281.

134   A. Dal Corso, S. Arosio, N. Arrighetti, P. Perego, L. Belvisi, L. Pignataro and C. Gennari, *Chem. Commun.*, 2021, **57**, 7778-7781.

135   M. A. Fanale, S. M. Horwitz, A. Forero-Torres, N. L. Bartlett, R. H. Advani, B. Pro, R. W. Chen, A. Davies, T. Illidge, D. Huebner, D. A. Kennedy and A. R. Shustov, *J. Clin. Oncol.*, 2014, **32**, 3137-3143.

136   K. Roy and P. L. A. Popelier, *J. Phys. Org. Chem.*, 2009, **22**, 186-196.

137   I. Sharma and G. A. Kaminski, *J. Comput. Chem.*, 2012, **33**, 2388-2399.

138   G. Chuchani and A. Frohlich, *J. Chem. Soc.*, 1971, DOI: 10.1039/J29710001417, 1417-1420.

139   R. M. Beesley, C. K. Ingold and J. F. Thorpe, *J. Chem. Soc., Trans.*, 1915, **107**, 1080-1106.

140   J. M. Barbosa-Filho, M. R. Piuvezam, M. D. Moura, M. S. Silva, K. V. B. Lima, E. V. L. da-Cunha, I. M. Fechine and O. S. Takemura, *Rev. Bras. Farmacogn.*, 2006, **16**, 109-139.

141   K. Mohan, R. Jeyachandran and R. Deepa, *Ann. Phytomed.*, 2012, **1**, 46-53.

142   T. T. Cushnie, B. Cushnie and A. J. Lamb, *Int. J. Antimicrob. Agents*, 2014, **44**, 377-386.

143   H. Sayhan, S. G. Beyaz and A. Çeliktaş, *Intech Open*, 2017, 57-84.

144   J. P. Krise and R. Oliyai, *Prodrugs: Challenges and Rewards, Part 1*, Springer, 2007, 801-831.

145   V. Abet, F. Filace, J. Recio, J. Alvarez-Builla and C. Burgos, *Eur. J. Med. Chem.*, 2017, **127**, 810-827.

146   E. Procházková, P. Šimon, M. Straka, J. Filo, M. Májek, M. Cigáň and O. Baszczyňski, *Chem. Commun.*, 2021, **57**, 211-214.

147   M. Mulliez and R. Wolf, *Bull. Soc. Chim. Fr.*, 1986, 101-108.

148   Schrödinger Release 2022-1 ed., Schrödinger, LLC, New York, NY, 2022.

149   J. E. Kilpatrick, K. S. Pitzer and R. Spitzer, *J. Am. Chem. Soc.*, 1947, **69**, 2483-2488.

150   C. Altona, H. J. Geise and C. Romers, *Tetrahedron*, 1968, **24**, 13-32.

151   C. Altona and M. Sundaralingam, *J. Am. Chem. Soc.*, 1972, **94**, 8205-8212.

152   N. S. Pagadala, K. Syed and J. Tuszynski, *Biophys. Rev.*, 2017, **9**, 91-102.

153   J. G. P. Wicker and R. I. Cooper, *J. Chem. Inf. Model.*, 2016, **56**, 2347-2352.

154   T. Montag and C. M. Thiele, *Chem. Eur. J.*, 2013, **19**, 2271-2274.

155   M. Tichotová, A. Ešnerová, L. Tučková, L. Bednárová, I. Císařová, O. Baszczyňski and E. Procházková, *J. Magn. Reson.*, 2022, **336**, 107149.

156   V. M. Marathias, G. J. Tawa, I. Goljer and A. C. Bach, *Chirality*, 2007, **19**, 741-750.

157   A. Marx, V. Schmidts and C. M. Thiele, *Magn. Reson. Chem.*, 2009, **47**, 734-740.

158   L. Vrzal, M. Kratochvilova-Simanova, T. Landovsky, K. Polivkova, J. Budka, H. Dvorakova and P. Lhotak, *Org. Biomol. Chem.*, 2015, **13**, 9610–9618.

159   T. Landovský, M. Tichotová, L. Vrzal, J. Budka, V. Eigner, H. Dvořáková and P. Lhoták, *Tetrahedron*, 2018, **74**, 902-907.

160   Schrödinger Release 2022-1: Glide, Schrödinger, LLC, New York, NY, 2021.

161   P. Bouř and T. A. Keiderling, *J. Chem. Phys.*, 2002, **117**, 4126-4132.

162   L.-J. Riwar, N. Trapp, B. Kuhn and F. Diederich, *Angew. Chem. Int. Ed.*, 2017, **56**, 11252-11257.

163   P. Bouř, *Collect. Czech. Chem. Commun.*, 2005, **70**, 1315-1340.

164   M. Di Pietro, P. Tzvetkova, T. Gloge, U. Sternberg and B. Luy, *Liq. Cryst.*, 2020, **47**, 1-15.

165   P. Tzvetkova, U. Sternberg, T. Gloge, A. Vazquez and B. Luy, *Chem. Sci.*, 2019, **10**, 8774-8791.

166   U. Sternberg, P. Tzvetkova and C. Muhle-Goll, *Phys. Chem. Chem. Phys.*, 2020, **22**, 17375-17384.

167   C. Farès, J. Lingnau, C. Wirtz and U. Sternberg, *Molecules*, 2019, **24**, 4417.

168   U. Sternberg and C. Farès, *Phys. Chem. Chem. Phys.* , 2022, **24**, 9608-9618.

169   C. M. Thiele, *Eur. J. Org. Chem.*, 2008, **2008**, 5673-5685.

7. Appendix

List of Papers

1. T. Landovský, **M. Tichotová**, L. Vrzal, J. Budka, V. Eigner, H. Dvořáková and P. Lhoták,* *Structure Elucidation of Phenoxathiin-Based Thiacalix[4]arene Conformations Using NOE and RDC Data*, Tetrahedron, *2018*, **74**, 902–907

2. T. Landovský, V. Eigner, M. Babor, **M. Tichotová**, H. Dvořáková and P. Lhoták,* *Regioselective S_NAr Reaction of the Phenoxathiin-Based Thiacalixarene as a Route to a Novel Macrocyclic Skeleton*, Chem. Commun., *2020*, **56**, 78–81

3. P. Šimon,[+] **M. Tichotová**,[+] M. García Gallardo, E. Procházková* and O. Baszczyňski,* *Phosphate-Based Self-Immolative Linkers for Tuneable Double Cargo Release*, Chem. Eur. J., *2021*, **27**, 12763–12775

4. M. Đud,[+] **M. Tichotová**,[+] E. Procházková* and O. Baszczyňski,* *Phosphate-Based Self-Immolative Linkers for the Delivery of Amine-Containing Drugs,* Molecules, *2021*, **26**, 5160

5. **M. Tichotová**, A. Ešnerová, L. Tučková, L. Bednárová, I. Císařová, O. Baszczyňski and E. Procházková,* *^{31}P NMR Parameters May Facilitate the Stereochemical Analysis of Phosphorus-Containing Compounds*, J. Magn. Reson., *2022*, **336**, 107149

6. A. Hruzíková,[+] L. Mužíková Čechová,[+] S. Štěpánová, L. Tučková, **M. Tichotová**, A. Růžička, V. Kašička and E. Procházková,* *A Study of Azopyrimidine Photoswitches and their Interactions with Cyclodextrins: When the Guest Governs the Type of Accommodation at the Host*, Dyes Pigm., *2023*, **212**, 111099

7. **M. Tichotová**, T. Landovský, J. Lang, S. Jeziorowski, V. Schmidts, M. Kohout, M. Babor, P. Lhoták,* C. M. Thiele* and H. Dvořáková*, *Enantiodiscrimination of Inherently Chiral Thiacalixarenes by Residual Dipolar Couplings*, J. Org. Chem., *2023*, DOI: 10.1021/acs.joc.2c02594

8. **M. Christou Tichotová**, L. Tučková, H. Kocek, A. Růžička, M. Straka and E. Procházková,* *Exploring the Impact of Alignment Media on RDC Analysis of Phosphorus-Containing Compounds: A Molecular Docking Approach*, Phys. Chem. Chem. Phys., *2024*, **26**, 2016–2024

9. U. Sternberg, **M. Christou Tichotová**, L. Tučková, A. Ešnerová, J. Hanus, O. Baszczyňski and E. Procházková,* *Extending Molecular Dynamics with Dipolar NMR Tensors as Constraints to Chiral Phosphorus Compounds*, Phys. Chem. Chem. Phys., *2024, under revision*

[+]these authors contributed equally

*corresponding author

List of Conferences

Lectures

1. *Phosphate-Based Self-Immolative Systems for Drug Delivery*

    M. Tichotová, P. Šimon, M. G. Gallardo, O. Baszczyňski, E. Procházková

    35th Central European NMR Meeting 2022, Valtice, Czech Republic

2. *^{31}P NMR with in situ Irradiation for Study of Self-Immolation*

    M. Tichotová, P. Šimon, M. G. Gallardo, O. Baszczyňski, E. Procházková

    Magnetic Moments in Central Europe 2022, Primošten, Croatia

3. *Structural Analysis Using ^{31}P NMR Spectroscopy Parameters*

    M. Tichotová, A. Ešnerová, L. Tučková, L. Bednárová, I. Císařová, O. Baszczyňski, U. Sternberg, E. Procházková

    36th Central European NMR Meeting 2023, Valtice, Czech Republic

4. *Exploring the Impact of Alignment Media on RDC Analysis of ^{31}P-Based Compounds*

    M. Christou, L. Tučková, H. Kocek, A. Růžička, M. Straka, E. Procházková

    37th Central European NMR Meeting 2024, Valtice, Czech Republic

Posters

1. *Phosphate-Containing Self-Immolative Linkers for Double Cargo Release*

    M. Tichotová, P. Šimon, M. G. Gallardo, O. Baszczyňski, E. Procházková

    73. Zjazd chemikov 2021, Vysoké Tatry, Slovakia

2. *Phosphate-Containing Self-Immolative Linkers for Double Cargo Release*

    M. Tichotová, P. Šimon, M. G. Gallardo, O. Baszczyňski, E. Procházková

    SMASH – Small Molecule NMR Conference 2021, online

3. *Phosphate Self-Immolative Linkers for Double Cargo Release*

    M. Tichotová, P. Šimon, M. G. Gallardo, O. Baszczyňski, E. Procházková

    55th Advances in Organic, Bioorganic and Pharmaceutical Chemistry 2021, Špindlerův Mlýn, Czech Republic

4. *^{31}P NMR Spectroscopy Parameters for a Structural Analysis*

    M. Tichotová, A. Ešnerová, L. Tučková, L. Bednárová, I. Císařová, O. Baszczyňski, E. Procházková

    28th International Conference on Organometallic Chemistry 2022, Prague, Czech Republic

5. *^{31}P NMR Spectroscopy Parameters for Structural Analysis*

    M. Tichotová, A. Ešnerová, L. Tučková, L. Bednárová, I. Císařová, O. Baszczyňski, E. Procházková

    55th Advances in Organic, Bioorganic and Pharmaceutical Chemistry 2022, Špindlerův Mlýn, Czech Republic

6. *^{31}P NMR Spectroscopy Parameters for Structural Analysis*

    M. Tichotová, A. Ešnerová, L. Tučková, L. Bednárová, I. Císařová, O. Baszczyňski, U. Sternberg, E. Procházková

    Experimental Nuclear Magnetic Resonance Conference 2023, Pacific Grove, California, USA

7. ^{31}P *NMR Spectroscopy Parameters for Structural Analysis*

   M. Tichotová, A. Ešnerová, L. Tučková, L. Bednárová, I. Císařová, O. Baszczyňski, E. Procházková

   Central European NMR Symposium and Bruker Users Meeting 2023, Prague, Czech Republic

8. ^{31}P *NMR Spectroscopy Parameters for Structural Analysis*

   M. Tichotová, A. Ešnerová, L. Tučková, L. Bednárová, I. Císařová, O. Baszczyňski, U. Sternberg, E. Procházková

   SMASH – Small Molecule NMR Conference 2023, Baveno, Italy

I.

P. Šimon, **M. Tichotová**, M. García Gallardo, E. Procházková and O. Baszczyňski:

Phosphate-Based Self-Immolative Linkers for Tuneable Double Cargo Release

Phosphate-Based Self-Immolative Linkers for Tuneable Double Cargo Release

Petr Šimon+,[a] Markéta Tichotová+,[a, b] María García Gallardo,[a] Eliška Procházková,*[b] and Ondřej Baszczyňski*[a, b]

We dedicate this paper to the memory of Dr. John C. Martin (1951–2021), head of Gilead Sciences, to commemorate his achievements in the development of anti-HIV drugs (acyclic nucleoside phosphonates, "Holy Trinity": A. Holý, E. De Clercq, J. C. Martin).

Abstract: Phosphorus-based self-immolative (SI) linkers offer a wide range of applications, such as smart materials and drug-delivery systems. Phosphorus SI linkers are ideal candidates for double-cargo delivery platforms because they have a higher valency than carbon. A series of substituted phosphate linkers was designed for releasing two phenolic cargos through SI followed by chemical hydrolysis. Suitable modifications of the lactate spacer increased the cargo release rate significantly, from 1 day to 2 hours or 5 minutes, as shown for linkers containing p-fluoro phenol. In turn, double cargo linkers bearing p-methyl phenol released their cargo more slowly (4 days, 4 hours, and 15 minutes) than their p-fluoro analogues. The α-hydroxyisobutyrate linker released both cargos in 25 minutes. Our study expands the current portfolio of SI constructs by providing a double cargo delivery option, which is crucial to develop universal SI platforms.

Introduction

Self-immolative (SI) linkers are chemical constructs subjected to irreversible fragmentation triggered by external stimuli.[1] After external activation, SI linkers disassemble (e.g., by cyclisation or electronic cascade), thereby releasing a leaving group (cargo). Triggered cargo release can be used in drug delivery[2] (prodrugs,[3] antibody-drug conjugates,[4] and chemosensors[5]), smart materials[6] (stimuli-responsive SI dendrimers[7] and polymers[8]), or in vivo cell labelling,[9] thus highlighting the wide range of applications of SI linkers. In turn, different structural motifs, including carbamate[3a,10] or phosphate,[11] can be used to attach chemically variable cargos to the SI linker.

Phosphorus-based SI linkers can offer higher versatility than their carbamate analogues because they allow us to attach an additional substituent (a second cargo), which significantly broadens their potential applications. For example, a double cargo linker can be used to simultaneously tether a drug (warhead, first cargo) and a reporter molecule (chromogenic molecule, second cargo).[12] Systems combining double (multi) cargo release have been studied[13] to pursue effective cancer chemotherapy using complementary drug combination,[14] antibiotic drug resistance[15] or novel fluorescent reporters/signal amplification chemosensors.[5a,16] In addition, phosphorus linkers enable us to monitor the reaction pathway by ^{31}P NMR spectroscopy in real time,[17] providing detailed structural and kinetic information, and to even detect the cyclic intermediates in some cases,[18] confirming cargo release by self-immolation, more specifically intramolecular cyclisation, and not by chemical hydrolysis.

Self-immolation, and consequently the rate of the cargo release reaction, can be tuned up by varying substituents on the phosphorus atom and is mainly directed by two effects: 1) the Thorpe-Ingold[19] effect and 2) the pK_a effect of the leaving groups. The Thorpe-Ingold effect refers to the spacer – sterically demanding substituents in the α-position accelerate intramolecular cyclisation.[20] Conversely, the pK_a effect is related to the cargo – the more acidic the cargo is (lower pK_a), the faster the cargo will be released upon external activation.[18] Fast cargo release systems are usually designed for biological applications. However, slow SI may also be advantageous when constructing slowly reacting (decomposing) polymers[8,21] or developing prodrugs with delayed cargo release (e.g., antibiotics).[22] Notwithstanding these advances and the potential of SI linkers, no phosphate-based, double cargo release system has been reported yet.

Considering the above, we designed a novel set of phosphate linkers bearing a lactate spacer (Figure 1) to attach two cargos (double cargo option). We also fine-tuned the release of both cargos by introducing a suitable modification in the lactate spacer and by adjusting the pK_a of the leaving groups. SI was triggered by UV light (365 nm) and monitored by ^{31}P NMR spectroscopy. The structures of intermediates and products were determined in situ by combining ^{13}C and [31]P

[a] Dr. P. Šimon,+ M. Tichotová,+ M. García Gallardo, Dr. O. Baszczyňski
Faculty of Science, Charles University
Prague 128 43 (Czech Republic)
E-mail: baszczyo@natur.cuni.cz

[b] M. Tichotová,+ Dr. E. Procházková, Dr. O. Baszczyňski
Institute of Organic Chemistry and Biochemistry
The Czech Academy of Sciences
Prague 166 10 (Czech Republic)
E-mail: prochazkova@uochb.cas.cz

[+] These authors contributed equally to this work.

Supporting information for this article is available on the WWW under
https://doi.org/10.1002/chem.202101805

Chemistry–A European Journal

Full Paper
doi.org/10.1002/chem.202101805

**Chemistry
Europe**
European Chemical
Societies Publishing
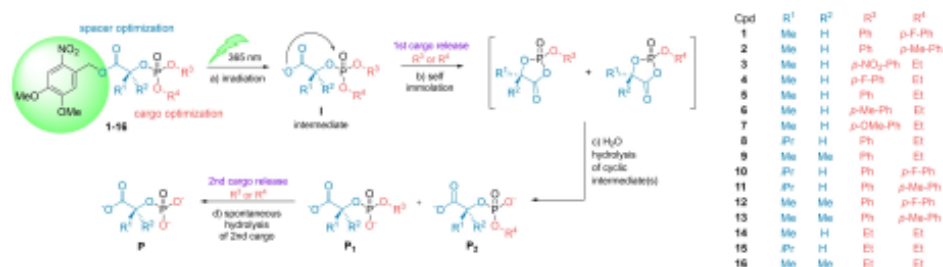
Figure 1. Chemical structure of phosphate-based linkers 1–16 and the proposed mechanism of self-immolation. a) Photoactivation, b) intramolecular cyclisation leading to the 1st cargo release; c) 2nd cargo release (through chemical hydrolysis).

NMR spectra. Our findings indicate that SI linkers with controlled double cargo release may be applied to develop novel combination therapies and smart materials.

Results and Discussion

Double cargo system: Proof of concept

Our initial strategy was to prepare the two model double cargo linkers with a lactate spacer as proof of concept. Compounds 1 and 2 were prepared by phosphorylation[23] of 22 using the corresponding in situ-generated diphenyl phosphorochloridates 18 and 19 (Scheme 1) in dichloromethane at room temperature. Photocleavable dimethoxy nitrobenzyl (DMNB)[24] ester 22 was prepared from DMNB-alcohol 20 and the lactic acid 21 via acid-catalysed esterification in refluxing toluene. Phosphorochlor-

idates 18 and 19 were synthesised from phenyl dichloridate 17 and the corresponding phenol in toluene at room temperature (Scheme 1), and 1 and 2 were synthesised using N-methyl-imidazole (NMI)[25] as a base/catalyst (Scheme 1c). Other bases, such as triethylamine (TEA), pyridine, and 4-(dimethylamino)pyridine (DMAP), were also tested during reaction optimisation, albeit without success.[26] The reaction progress was monitored by ^{31}P NMR spectroscopy (for NMR chemical shifts in CDCl$_3$, Table S9 in the Supporting Information).

We performed irradiation NMR experiments on 1 and 2, in a previously optimised mixture of cacodylate buffer/DMSO (1:1, v/v). Indeed, both compounds 1 and 2 cyclised upon UV light irradiation, releasing the cargos overnight (Figure 2). Compound 1, bearing the electron-withdrawing (EWG) p-F substituent, provided two ^{31}P NMR signals at $\delta_p = -13.02$ and -13.10 ppm, corresponding to two diastereoisomers resulting from a stereogenic centre on the phosphorus atom, and carbon

Scheme 1. Synthesis of the two model double cargo linkers 1 and 2 prepared for a proof of concept study. a) "Corresponding phenol", TEA, toluene, 25 °C, 12 h; b) p-toluenesulfonic acid, toluene, reflux, 16 h; c) diphenyl chlorophosphate 18 (for 1) or 19 (for 2) NMI, CH$_2$Cl$_2$, 25 °C, 12 h.
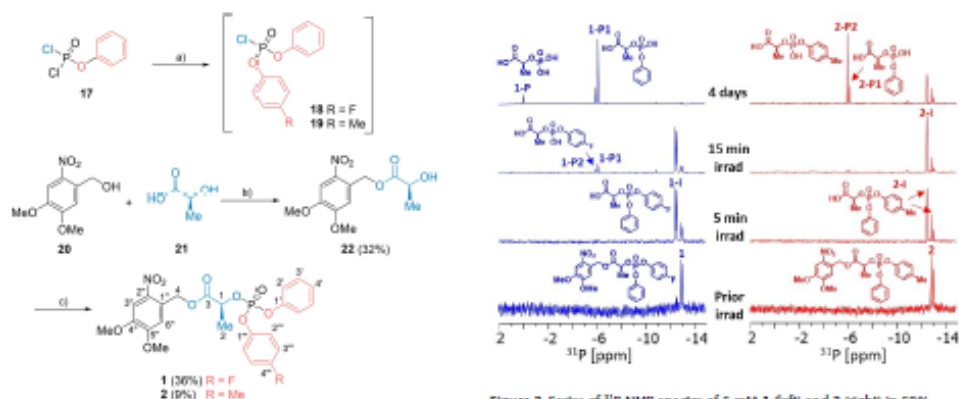
Figure 2. Series of ^{31}P NMR spectra of 5 mM 1 (left) and 2 (right) in 50% CACO/[D$_6$]DMSO recorded before and after UV light irradiation (365 nm) at room temperature. In both cases, the more acidic phenol (p-F-phenol in 1 and the unsubstituted phenol in 2) was released preferentially. The high signal-to-noise ratios of the first spectra (the prior irrad row) are caused by the lower solubility of the starting compounds in 50% CACO/[D$_6$]DMSO.

Chem. Eur. J. 2021, 27, 12763–12775   www.chemeurj.org

12764

© 2021 Wiley-VCH GmbH

Chemistry–A European Journal

Full Paper
doi.org/10.1002/chem.202101805

Chemistry
Europe
European Chemical
Societies Publishing

In the lactate spacer. Upon UV light irradiation, DMNB was cleaved off, and intermediate 1-I ($\delta_P = -12.55$ and -12.64 ppm) was formed in 5 minutes (full conversion of 1 to 1-I in 15 minutes). The activated linker 1-I slowly cyclised and afforded mono-phenyl products 1-P1 ($\delta_P = -6.17$ ppm) and 1-P2 ($\delta_P = -5.98$ ppm) overnight. The more acidic p-F-phenyl substituent (pK$_a$ 9.95[27]) was released by SI first (1-P1), while the phenyl substituent (with higher pK$_a$ 9.98[28]) remained mostly unscathed (1-P2). A trace amount of the fully hydrolysed product P ($\delta_P = 0.06$ ppm) was observed after 4 days. Compound 2, with the electron-donating (EDG) p-Me group ($\delta_P = -12.92$ and -13.09 ppm), also afforded the activated intermediate 2-I ($\delta_P = -12.57$ ppm), but its cyclisation proceeded significantly slower than that of 1-I. After 24 hours, 2-I was still the major component (Figure 2). The desired products 2-P1 ($\delta_P = -6.19$ ppm) and 2-P2 ($\delta_P = -6.02$ ppm) were identified, but the more acidic phenol (2-P2) was preferentially released.

Both linkers (1 and 2 differing in R$_4$, Figure 1) provided the same mono-phenyl product P1 (1-P1 and 2-P1), albeit on significantly different timescales (in 15 minutes and 24 hours, respectively). Phenol release can be controlled by a nature of the second cargo. The ability to attach cargos with different release rates might, hence, be useful for further applications, such as generating phosphorylated metabolites in biology experiments and drug delivery, among others.

The structures of the two intermediates (1-I and 2-I) and of products P1 and P2 were determined in situ by combining ^{13}C and ^{31}P NMR spectra. The key connectivity information was derived from the ^{13}C signal splitting caused by ^{13}C–^{31}P spin–spin interaction (J_{CP}), typically through two or three chemical bonds,

with $J_{CP} = 3–8$ Hz (Tables S3–S8). The ^{13}C–^{19}F spin–spin interactions ($J_{CF} = 4–244$ Hz) enabled us to easily identify the phenyl that was preferentially released from 1.

pK$_a$ effect on single-cargo release

To better understand the pK$_a$ effect on cargo release, we prepared a series of compounds 3–7 bearing only one releasable phenolic cargo differing in para substitution (Scheme 2). The lactate derivative 22 was treated with highly reactive ethyl dichloridate 23 and with one equivalent of TEA in toluene at room temperature to give the corresponding phosphorochloridate 24 (not isolated). Toluene was removed at low pressure, and crude 24 was subsequently used in reactions with the corresponding phenols in CH$_2$Cl$_2$, in the presence of TEA, to give 3–7. The reaction progress was monitored by ^{31}P NMR spectroscopy (for NMR chemical shifts in CDCl$_3$, Table S10).

Compounds 3–7 released the phenolic cargo successfully at various rates, and the same final product P2 ($\delta_P = -1.14$ ppm) was detected in all cases (Figure 3). In 3 (EWG NO$_2$ group), the SI was so fast that 3-I was not detected, and p-NO$_2$-phenol was released in an hour. Moreover, we observed spontaneous partial hydrolysis of 3 overnight, which released p-NO$_2$-phenol without photoactivation; therefore, no SI could proceed (hydrolytic product hP, Figure 3, left).

Such a hydrolytic decomposition of 3 may be easily misinterpreted for SI when using common optical methods.[18] SI was slower in 4 than in 3, the intermediate 4-I was detected, and the corresponding phenol was released in 24 h. The relative
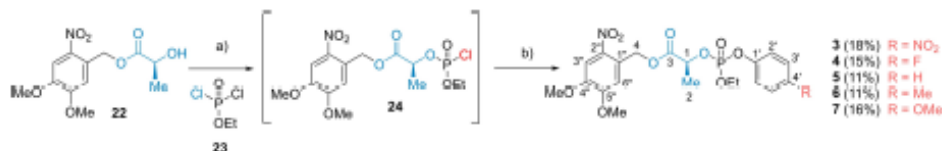
Scheme 2. Synthesis of lactate linkers 3–7. a) TEA, toluene, 25 °C, 12 h; b) "corresponding phenol", TEA, CH$_2$Cl$_2$, 25 °C, 12 h.
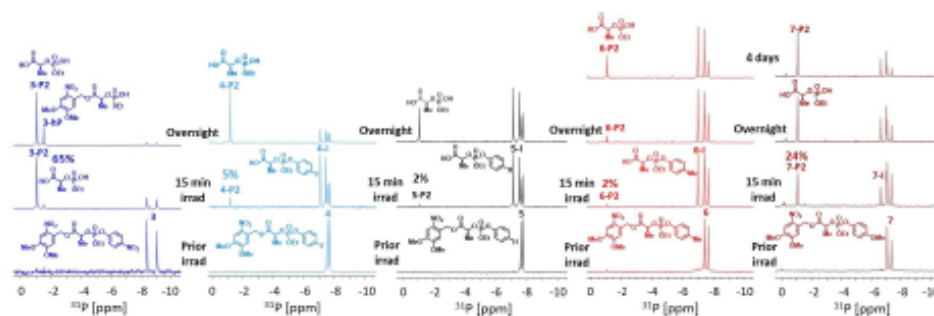
Figure 3. Series of ^{31}P NMR spectra of compounds 3–7 (5 mM, 50% CACO/[D$_6$]DMSO) recorded upon UV light irradiation (365 nm) at room temperature.

Chemistry–A European Journal

Full Paper
doi.org/10.1002/chem.202101805

Chemistry
Europe
European Chemical
Societies Publishing

concentration of products **3-P2** and **4-P2** was significantly higher than that of their unsubstituted phenol counterpart **5**. The SI reaction in **3**, **4** and **5** matches the corresponding pK_a values of p-NO$_2$, p-F, and p-H phenols (7.15,[28] 9.95, and 9.98, respectively). The lower the pK_a of phenol is, the faster SI will be. Conversely linkers bearing EDG phenols with p-Me (**6**) and p-OMe (**7**) released the corresponding phenol in 96 and 24 hours, respectively. The higher pK_a values of EDG-phenols in **6** and **7** (10.14[28] and 10.55,[29] respectively) resulted in slower cargo release in **6**; as such, only traces of **6-P2** were detected overnight. Yet, in contrast to **6**, linker **7** did not fully match its pK_a value and led to a faster cargo release than **5**. The 15-minute trace in Figure 3 shows the relative concentration of the final product P2 of **3–7** as follows: 65, 5, 2, 2 and 24%.

We monitored the cyclisation of **3–7** by ^{31}P NMR spectra with *in situ* irradiation (details in the Supporting Information) and extracted the concentration profiles of intermediates within 60 minutes of UV irradiation. Based on our previous study,[20] the photoactivation is similar among **3–7**, which bear the same lactate spacer. Thus, alterations in intermediate concentrations represent differences in cyclisation rates. For example, in 30 minutes, we obtain 0% of **3-I**, meaning that cyclisation is fast and that photoactivation is the rate-limiting step. Conversely, linkers **4–7** provided 32% **4-I**, 48% **5-I** and 55% **6-I**, and 18% **7-I** in 30 minutes (kinetic curves in Figure S1). The SI rates in **3–7** did not fully match the pK_a values of the leaving groups, showing the following trend: **6** (p-Me) < **5** (p-H) < **4** (p-F) < **7** (p-OMe) < **3** (p-NO$_2$). Most likely, the pK_a value is not the only parameter that affects SI. We may, nevertheless, speculate that a resonance-based change of electron density on phosphorus is mediated by p-OMe (free electron pair of oxygen), which supports SI.

Spacer optimisation

To accelerate cargo release, we used the Thorpe-Ingold effect[30] and modified the lactate spacer responsible for cyclisation in **5** by introducing either a sterically demanding or an additional α-substituent. Thus, we prepared model α-hydroxyisovalerate and α-hydroxyisobutyrate linkers (**8** and **9**, respectively). For this purpose, photocleavable (DMNB)-esters **28** and **29** were synthesised from DMNB-alcohol **20** and the corresponding carboxylic acids, α-hydroxyisovaleric **26** and α-hydroxyisobutyric **27**, through an acid-catalysed esterification in refluxing toluene using a condenser trap to remove the reaction water (Scheme 3). However, the synthetic approach used for compounds **3–7** failed in the derivatives of α-hydroxyisovalerate **8** and α-hydroxyisobutyrate **9**, most likely due to the degradation of the starting materials **26** and **27** (dehydration of branched acid esters **26** and **27**). Therefore, we came up with an alternative synthetic approach, similar to **1** and **2**, generating the phosphorylating agent **25** *in situ* to phosphorylate compounds **28** and **29** and thus yielding **8** and **9** (Scheme 3).

Monochloridate **25** was prepared in toluene using TEA. However, the phosphorylation of the sterically hindered alcohols **28** and **29** was challenging. Several conditions (solvents,
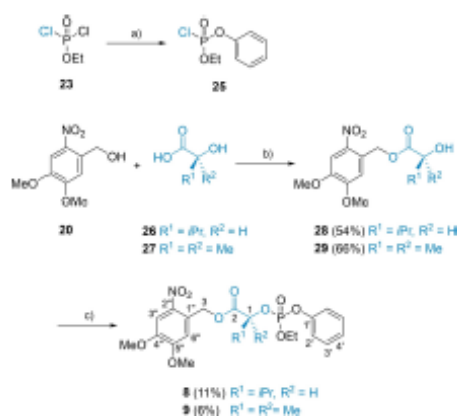
Scheme 3. Synthesis of linkers 8 and 9 with branched substituents in the α position. a) Phenol, TEA, toluene, 25 °C, 12 h; b) p-toluenesulfonic acid, toluene, reflux, 16 h; c) for 8: compound 25, DMAP, CH$_2$Cl$_2$, 25 °C, 12 h; for 9: compound 25, NMI, CH$_2$Cl$_2$, 25 °C, 12 h.

bases and temperatures) were tested, but only limited combinations gave the desired products, as monitored by ^{31}P NMR (Table S11). The secondary alcohol **28** yielded the desired product **8** only when using the DMAP base, and the tertiary alcohol **29** gave product **9** when using NMI, both of which in CH$_2$Cl$_2$.

Surprisingly, despite the increase in steric demands in the α position, SI was not significantly better in **8** than in **5** (Figure 4). Although traces of **8-P2** (δ_P −0.60 ppm) were detected after 15 minutes, a comparable amount of P2 from both analogues (**5** and **8**) was detected within 24 h. This finding contrasts with our previous observation in the phosphoramidate linker series,[20] wherein the increase in the α-substitution effect (Me versus iPr) markedly increased the cargo release rate. Conversely, linker **9** (δ_P −10.99 ppm) increased the cargo release considerably, with traces of **9-P2** (δ_P −4.22 ppm) being detected within 5 minutes. After 15 minutes, **9-P2** was the main component in the reaction mixture. This trend is also clearly visible in the concentration profiles of intermediates **5-I**, **8-I** and **9-I** (Figure S1), where **9-I** remains at approximately 20% (in 20–60 minutes), which means that **9-I** cyclises more quickly than **5-I** and **8-I**. In contrast, the concentration of **5-I** and **8-I** increases from 40 to 70% (in 20–60 minutes), indicating significantly slower cyclisation and, thus, slower cargo release.

Double cargo linkers with a tuneable release rate

Our structure-activity relationship study, including spacer and cargo modifications, provided us with deeper insights into SI and its limitations, encouraging us to design other double-cargo systems with a wider range of release rates as the basis for linkers with timed, double-cargo sequential release.
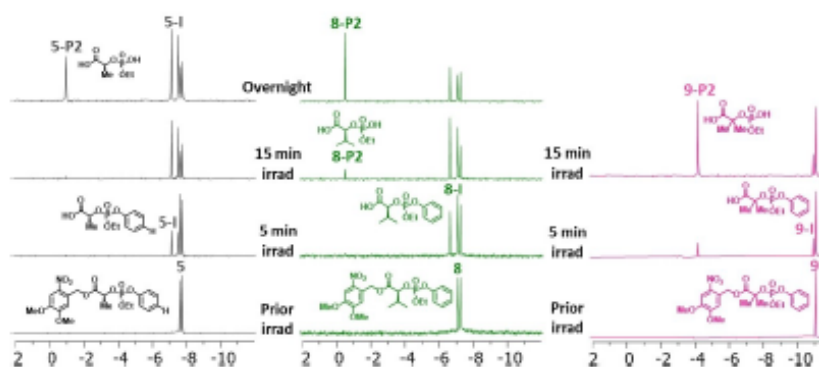
Chemistry–A European Journal

Full Paper
doi.org/10.1002/chem.202101805

Chemistry Europe
European Chemical
Societies Publishing

Figure 4. Series of ^{31}P NMR spectra of linkers 5, 8 and 9 (5 mM, 50% CACO/[D$_6$]DMSO) recorded upon UV light irradiation (365 nm) at room temperature.

The third class of linkers, **10–13**, containing diphenyl substituted phosphorus, was prepared similarly to **1** and **2**. *In situ*-generated phosphorochloridates **18** and **19** (Scheme 1) were directly used to phosphorylate **28** and **29** (Scheme 4), but sterically hindered alcohols **28** and **29** showed low reactivity to **18** and **19**. Thus, the reaction conditions required optimisation. For α-hydroxyisovalerate **28**, using DMAP in CH$_2$Cl$_2$ was the most effective approach. In contrast, for α-hydroxyisobutyrate **29**, using one equivalent of TEA in CH$_2$Cl$_2$ with DMAP catalysis led to the highest yields (Scheme 4). Tertiary alcohol **29** did not react under NMI conditions (as in **9**), nor did DMAP (without TEA) or TEA (without DMAP). This difference in reactivity is likely caused by the low reactivity of diphenyl monochlorides **18** and **19** combined with the bulky alcohols **28** and **29**. Hence, synthetic access to asymmetric diphenol-phosphates with a bulky alcohol spacer should be optimised for particular substrates. In this study, we monitored the progress of the reaction by ^{31}P NMR spectroscopy (for NMR chemical shifts in CDCl$_3$, see Table S12; for the synthetic note, see p. S23). All double-cargo linkers **10–13** displayed the desired properties, covering a wide range of cargo release rates – from minutes to days – and enabling sequential cargo release (Figure 5).
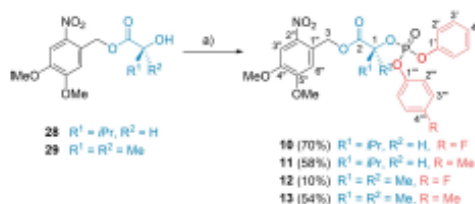
Modifying the lactate spacer increased SI significantly, from 24 hours in **1** and **2** to 2 hours in α-hydroxyisovalerate analogues **10** and **11** or to 15 minutes in their α-hydroxybutyrate counterparts **12** and **13**, as shown in Figures 2 and 5. The SI of **10–13** matches the trend found in single-cargo linkers **5**, **8** and **9** (Figure 4) where butyrate spacer showed faster cyclisation (also supported by the concentration profiles of the intermediates, Figure S1). Moreover, the relative concentrations of **12-I** and **13-I** are approximately 15 and 35%, which proves the faster SI and cargo release of *p*-F linker **12**. We also noticed that the cargo release rates of **10** and **12** (EWG cargos, *p*-F phenyl group) are similar to those of **1** and **11** and that the cargo release rate of **13** (EDG cargos, *p*-Me phenyl moiety) is similar to that of **2**. In general, linkers **1**, **10**, and **12** with EWG cargo yielded a significantly higher amount of P1/P2 than their EDG cargo counterparts **2**, **11**, and **13**, respectively, as shown in Figures 2 and 5 (see the 15-min row). Moreover, the α-hydroxyisovalerate analogues **10** and **11** were measured in 25% CACO/DMSO because of their low solubility in the 50% solvent system (the solvent effect is shown in Figure S2, and is in line with our previous study).

Interestingly, in **12**, both phenolic cargos were released within 25 minutes, exclusively yielding the final product **12-P**, whereas compound **13** gave a mixture of **13-P2** and **13-P** (= **12-P**) in 3 hours, showing similar characteristics to **11** but a faster release rate. Importantly, **12** and **13** are stable in 50% CACO/DMSO mixture at room temperature for seven days (Figure S3), which is enough time for most biological applications.

Negative controls

We examined three linkers, **14–16**, differing in the spacer (lactate, α-hydroxyisovalerate and α-hydroxybutyrate, respectively) and bearing ethyl moieties instead of phenyl groups serving as negative controls. We expected that the photo-activation step would result in the formation of intermediate I

Scheme 4. Synthesis of branched double-cargo linkers **10–13**. a) For **10** and **11**: "diaryl-chlorophosphate **18** or **19**", DMAP, CH$_2$Cl$_2$, 25 °C, 12 h; for **12** and **13**: "diaryl-chlorophosphate **18** and **19**", respectively, TEA, DMAP (cat.), CH$_2$Cl$_2$, 25 °C, 12 h.

28 R^1 = *i*Pr, R^2 = H
29 R^1 = R^2 = Me

10 (70%) R^1 = *i*Pr, R^2 = H, R = F
11 (58%) R^1 = *i*Pr, R^2 = H, R = Me
12 (10%) R^1 = R^2 = Me, R = F
13 (54%) R^1 = R^2 = Me, R = Me

Chemistry–A European Journal

Full Paper
doi.org/10.1002/chem.202101805

Chemistry
Europe
European Chemical
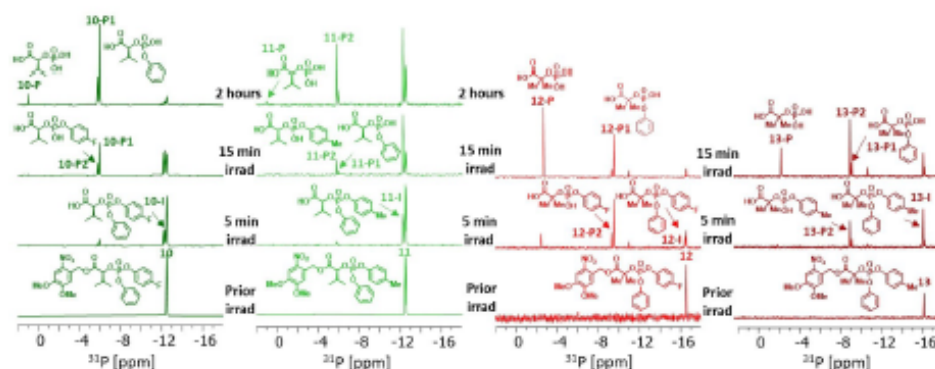Societies Publishing

Figure 5. Series of ^{31}P NMR spectra of compounds 10–13 (5 mM, 25% (10, 11) or 50% (12, 13) CACO/[D$_6$]DMSO) recorded upon UV light irradiation (365 nm) at room temperature. The more acidic phenolic cargo is preferentially released; the α-hydroxybutyrate linkers 12 and 13 released the cargo within minutes, whereas the α-hydroxyisovalerate analogues 10 and 11 released the cargo within hours.

without any further change (no cargo release). As a second negative control, we included three derivatives, 32–34, bearing no SI spacer. All negative controls, 14–16 and 32–34, confirmed the importance of a suitable cargo (appropriate leaving group) and a spacer, which is responsible for SI (Scheme 5). Without them, no SI cascade is possible, as shown in Figure S5 and Section 2 in the Supporting Information.

Conclusion

We have developed a novel class of phosphate-based SI linkers with a tuneable double-cargo release option. The first cargo is released through SI; the second cargo is released through chemical hydrolysis. Suitable modifications of the lactate spacer increased the cargo release rate significantly, from 1 day to 2 hours or to 5 minutes, as shown for the linkers containing p-F phenol 1, 10, or 12, respectively. In turn, the double-cargo linkers 2, 11, and 13, bearing p-Me phenyl, released their cargo more slowly (4 days, 4 hours, and 15 minutes, respectively) than their p-F analogues. Our linkers provide a) the programmable release of the first cargo in 3 hours (10-P1, 11-P2), b) simultaneous release of both cargos within 25 minutes (12-P) or

c) sequential release of the first cargo within 15 minutes (13-P2) followed by release of the second cargo within 3 hours (13-P). Linkers 12 and 13, with the fastest cargo release rates, have shown satisfying stability in a 50% buffer/DMSO mixture for seven days at room temperature. Overall, we are now able to drive the sequential release of two cargos from minutes to days by controlling SI. Our systems also provide phosphorylated products P1, P2, and P over a wide range of timescales. These systems may thus find further applications, such as the development of new smart materials and multiple drug delivery or the generation of phosphorylated species. Ultimately, our study pioneers a novel route for the design of phosphorus-based, self-immolative systems for double-cargo release.

Experimental Section

General. All reagents were purchased from commercial suppliers and used as received. 4,5-Dimethoxy-2-nitrobenzyl alcohol was purchased from Fluorochem Ltd. (UK). All reactions were performed under an inert argon atmosphere. Thin layer chromatography (TLC) was performed on TLC aluminium sheets (silica-gel 60 F$_{254}$; Merck). Reaction progress was monitored using TLC and/or ^{31}P NMR spectroscopy in CDCl$_3$. Flash-column chromatography was performed on a Compact (ECOM s.r.o.) chromatography system using silica-gel or C$_{18}$ silica-gel 230–400 mesh, 60 Å (Merck KGaA, Germany). The final products were recovered by solvent evaporation. All products were viscous oils, semi-solids or noncrystalline solids. The reaction yields were not optimised.

Synthesis

4,5-Dimethoxy-2-nitrobenzyl L-lactate (22). 4,5-Dimethoxy-2-nitrobenzyl alcohol 20 (213 mg, 1.00 mmol, 1.00 equiv.) and L-lactic acid (90 mg, 1.00 mmol, 1.00 equiv.) were suspended in toluene (50 mL), subsequently adding p-toluenesulfonic acid (50 mg, 0.26 mmol, 0.26 equiv.) at 25 °C. The mixture was refluxed for 12 h, evaporated to dryness *in vacuo* with silica-gel, and the title compound was

Scheme 5. Negative controls: linkers 14–16 bearing no releasable ethanol cargo (left), and linkers 32–34 with no spacer responsible for SI (right).

14 (13%) R^1 = Me, R^2 = H
15 (17%) R^1 = iPr, R^2 = H
16 (11%) R^1 = Me, R^2 = Me

32 (21%) R^3 = Et, R^4 = Et
33 (17%) R^3 = Ph, R^4 = Et
34 (11%) R^3 = Ph, R^4 = p-F-Ph

Chemistry–A European Journal

Full Paper
doi.org/10.1002/chem.202101805

Chemistry
Europe
European Chemical
Societies Publishing

isolated by normal-phase flash chromatography (CH_2Cl_2/methanol gradient on silica gel) and followed by reversed-phase flash chromatography (water/acetonitrile gradient on C_{18} silica gel). Yield **22** (90 mg, 32%) of an off-white semisolid. ^1H NMR (400 MHz, [D_6] DMSO, 25 °C): δ = 7.71 (s, 1H, 3''), 7.21 (s, 1H, 6''), 5.55 (d, 1H, J_{OH-3} = 5.9, –OH), 5.46 (d, 1H, J_{GEM} = 14.1, 4a), 5.39 (d, 1H, J_{GEM} = 14.1, 4b), 4.25 (m, 1H, 1), 3.90 (s, 3H, 5''–O–CH_3), 3.87 (s, 3H, 4''–O–CH_3), 1.29 ppm (d, 3H, J_{2-1} = 6.8, 2). ^{13}C NMR (100 MHz, [D_6]DMSO, 25 °C): δ = 174.07 (3), 153.18 (5''), 147.98 (4''), 139.69 (2''), 126.16 (1''), 111.35 (6''), 108.28 (3''), 65.99 (1), 62.62 (4), 56.26 (5''–O–CH_3), 56.12 (4''–O–CH_3), 20.32 ppm (2). HRMS (ESI+) calculated for $C_{12}H_{15}O_7NNa$: 308.07407; found: $[M+Na]^+$ 308.07381.

4,5-dimethoxy-2-nitrobenzyl (S)-2-hydroxy-3-methylbutanoate (28). 4,5-Dimethoxy-2-nitrobenzyl alcohol **20** (2.55 g, 12.0 mmol, 1.20 equiv.) and *p*-toluenesulfonic acid (172 mg, 1.00 mmol, 0.10 equiv.) were suspended in toluene (100 mL), subsequently adding (S)-2-hydroxy-3-methylbutanoic acid (1.18 g, 10.0 mmol, 1.00 equiv.) after 45 min of azeotropic distillation. The mixture was refluxed overnight and washed with saturated $NaHCO_3$ (60 mL), water (60 mL), and saturated NaCl (60 mL). The organic phase was dried with Na_2SO_4 and evaporated to dryness *in vacuo* with silica-gel. The title compound was isolated by normal-phase flash chromatography (CH_2Cl_2/methanol gradient on silica-gel) followed by reversed-phase flash chromatography (water/methanol gradient on C_{18} silica-gel). Yield **28** (1.68 g, 54%) of a red-brown semisolid. ^1H NMR (400 MHz, CDCl$_3$, 25 °C): δ = 7.75 (s, 1H, 3''), 7.01 (s, 1H, 6''), 5.67 (d, J_{GEM} = 14.3, 3a), 5.59 (d, J_{GEM} = 14.3, 3b), 4.16 (d, 1H, $J_{1-CH(iPr)}$ = 3.5, 1), 4.00 (s, 3H, 5''–O–CH_3), 3.99 (s, 3H, 4''–O–CH_3), 2.14 (m, CH^{iPr}), 1.07 (d, 3H, $J_{CH_3(iPr)-CH(iPr)}$ = 7.0, CH$_3^{iPr}$), 0.92 ppm (d, 3H, $J_{CH_3(iPr)-CH(iPr)}$ = 7.0, CH$_3^{iPr}$). ^{13}C NMR (100 MHz, CDCl$_3$, 25 °C): δ = 174.38 (2), 153.45 (5''), 148.57 (4''), 140.20 (2''), 125.90 (1''), 110.89 (6''), 108.33 (3''), 75.16 (1), 64.19 (3), 56.42–56.46 (m, 4''–O–CH_3 and 5''–O–CH_3), 32.26 (CH^{iPr}), 18.78 and 16.03 ppm (CH$_3^{iPr}$). HRMS (ESI+) calculated for $C_{14}H_{19}O_7NNa$: 336.10537; found: $[M+Na]^+$ 336.10506.

4,5-Dimethoxy-2-nitrobenzyl 2-hydroxy-2-methylpropanoate (29). 4,5-Dimethoxy-2-nitrobenzyl alcohol (2.13 g, 10.0 mmol, 1.00 equiv.) and 2-hydroxy-2-methylpropanoic acid (1.04 g, 10.0 mmol, 1.00 equiv.) were suspended in toluene (100 mL), subsequently adding *p*-toluenesulfonic acid (200 mg, 1.04 mmol, 0.10 equiv.) at 25 °C. The mixture was refluxed for 12 h and evaporated to dryness *in vacuo* with silica-gel, isolating the title compound by normal-phase flash chromatography (CH_2Cl_2/methanol gradient on silica-gel) followed by reversed-phase flash chromatography (water/acetonitrile gradient on C_{18} silica-gel). Yield **29** (1.97 mg, 66%) of a greyish semisolid. ^1H NMR (400 MHz, CDCl$_3$, 25 °C): δ = 7.75 (s, 1H, 3''), 6.98 (s, 1H, 6''), 5.61 (s, 2H, 3), 3.99 and 3.98 (s, 6H, 5''–O–CH_3 and 4''–O–CH_3), 1.52 ppm (s, 6H, 1-(CH_3)$_2$). ^{13}C NMR (100 MHz, CDCl$_3$, 25 °C): δ = 176.69 (2), 153.49 and 148.42 (5'' and 4''), 139.96 (2''), 126.32 (1''), 110.16 (6''), 108.34 (3''), 72.19 (1), 64.34 (3), 56.42 and 56.38 (5''–O–CH_3 and 4''–O–CH_3), 27.21 ppm (1-(CH_3)$_2$). HRMS (ESI+) calculated for $C_{13}H_{17}O_7NNa$: 322.08972; found: $[M+Na]^+$ 322.08944.

4,5-Dimethoxy-2-nitrobenzyl (2S)-2-(((4-fluorophenoxy)(phenoxy) phosphoryl)oxy)propanoate (1). Phenyl dichlorophosphate **17** (60 μL, 0.40 mmol, 1.00 equiv.) was dissolved in toluene (3 mL), subsequently adding TEA (100 μL, 0.72 mmol, 1.80 equiv.) and 4-fluorophenol (45 mg, 40 mmol, 1.00 equiv.) at 25 °C. The mixture was stirred at 25 °C for 12 h and evaporated to dryness *in vacuo*. The solid residue was dissolved in CH_2Cl_2 (3 mL) and 4,5-dimethoxy-2-nitrobenzyl L-lactate **22** (110 mg, 0.39 mmol, 0.98 equiv.), subsequently adding NMI (35 μL, 0.44 mmol, 1.10 equiv.) at 25 °C. The mixture was stirred at 25 °C for 12 h and evaporated to dryness *in vacuo*. The title compound was isolated by normal-phase flash chromatography (n-hexane/ethyl acetate gradient on silica-gel) followed by reversed-phase flash chromatography (water/

acetonitrile gradient on C_{18} silica-gel). Yield **1** (76 mg, 36%) of a yellow oil. NOTE: approximately 1:1 mixture of diastereomers. ^1H NMR (400 MHz, CDCl$_3$, 25 °C): δ = 7.71 (s, 1H, 3''), 7.71 (s, 1H, 3''), 7.35 (m, 2H, 3'), 7.29 (m, 2H, 3'), 7.14–7.24 (m, 10H, 4', 2''', 2'), 7.02 (m, 2H, 3''), 6.97 (m, 2H, 3'''), 7.00–7.01 (m, 2H, 6''), 5.62 (dd, 1H, J_{GEM} = 14.9, J_{4b-c} = 0.6, 4a), 5.61 (dd, 1H, J_{GEM} = 14.7, J_{4b-c} = 0.6, 4a), 5.56 (dd, 1H, J_{GEM} = 14.7, J_{4b-c} = 0.6, 4b), 5.53 (dd, 1H, J_{GEM} = 14.9, J_{4b-c} = 0.6, 4b), 5.15–5.24 (m, 2H, 1), 3.96 (s, 3H, 4''–O–CH_3), 3.95 (s, 3H, 4''–O–CH_3), 3.94 (s, 3H, 5''–O–CH_3), 3.92 (s, 3H, 5''–O–CH_3), 1.58–1.64 ppm (m, 6H, 2). ^{13}C NMR (100 MHz, CDCl$_3$, 25 °C): δ = 169.47 (d, J_{3-P} = 5.0, 3), 169.38 (d, J_{3-P} = 4.7, 3), 160.13 (dd, $J_{4'-F}$ = 244.8, $J_{4'-P}$ = 1.1, 4'''), 160.08 (dd, $J_{4'-F}$ = 244.2, $J_{4'-P}$ = 1.1, 4'''), 153.89 and 153.88 (5''), 150.25 (d, $J_{1'-P}$ = 7.3, 1'), 150.16 (d, $J_{1'-P}$ = 7.3, 1'), 148.54 and 148.50 (4''), 146.15–146.45 (m, 1'''), 139.76–139.94 (m, 2''), 130.04 and 129.91 (3'), 126.33 and 126.24 (1''), 125.87 (d, $J_{4'-P}$ = 1.4, 4'), 125.78 (d, $J_{4'-P}$ = 1.5, 4'), 121.86 (d, $J_{2'-P}$ = 4.6, 2'''), 121.70 (d, $J_{2'-P}$ = 4.9, 2'''), 120.24 (d, $J_{2'-P}$ = 4.7, 2'), 120.20 (d, $J_{2'-P}$ = 5.4, 2'), 116.59 (d, $J_{3'-F}$ = 23.6, 3'''), 116.43 (d, $J_{3'-F}$ = 23.6, 3'''), 110.43 and 110.37 (6''), 108.31 and 108.30 (3''), 73.65 (d, J_{1-P} = 6.1, 1), 73.64 (d, J_{1-P} = 5.9, 1), 64.53 (4), 56.75 (5''–O–CH_3), 56.55 (4''–O–CH_3), 19.16 (d, J_{2-P} = 5.9, 2), 19.13 ppm (d, J_{2-P} = 6.2, 2). ^{31}P NMR (161 MHz, CDCl$_3$, 25 °C): δ = −12.33 ppm. HRMS (ESI+) calculated for $C_{24}H_{23}O_{10}NFNaP$: 558.09358; found: $[M+Na]^+$ 558.09332.

4,5-Dimethoxy-2-nitrobenzyl (2S)-2-(((4-methylphenoxy)(phenoxy) phosphoryl)oxy)propanoate (2). Phenyl dichlorophosphate **17** (60 μL, 0.40 mmol, 1.00 equiv.) was dissolved in toluene (3 mL), subsequently adding TEA (100 μL, 0.72 mmol, 1.80 equiv.) and 4-methylphenol (45 mg, 0.40 mmol, 1.00 equiv.) at 25 °C. The mixture was stirred at 25 °C for 12 h and evaporated to dryness *in vacuo*. The solid residue was dissolved in CH_2Cl_2 (3 mL) and 4,5-dimethoxy-2-nitrobenzyl L-lactate **22** (114 mg, 0.40 mmol, 1.02 equiv.), adding NMI (40 μL, 0.50 mmol, 1.26 equiv.) at 25 °C. The mixture was stirred at 25 °C for 12 h and evaporated to dryness *in vacuo*. The title compound was isolated by normal-phase flash chromatography (n-hexane/ethyl acetate gradient on silica-gel) followed by reversed-phase flash chromatography (water/methanol gradient on C_{18} silica-gel). Yield **2** (20 mg, 9%) of a yellowish oil. NOTE: approximately 1:1 mixture of diastereomers. ^1H NMR (400 MHz, CDCl$_3$, 25 °C): δ = 7.72 (s, 1H, 3''), 7.71 (s, 1H, 3''), 7.31–7.37 (m, 2H, 3'), 7.26–7.31 (m, 2H, 3'), 7.19–7.24 (m, 4H, 2'''or 2'), 7.12–7.19 (m, 2H, 4'), 7.09–7.12 (m, 2H, 3'''), 7.06–7.09 (m, 6H, 3''', 2' or 2'''), 7.02–7.04 (m, 2H, 6''), 5.51–5.64 (m, 4H, 4), 5.16–5.25 (m, 2H, 1), 3.95–3.96 (m, 6H, 4''–O–CH_3), 3.92 (s, 3H, 5''–O–CH_3), 3.91 (s, 3H, 5''–O–CH_3), 2.32 (s, 3H, 4'''-CH_3), 2.28 (s, 3H, 4'''-CH_3), 1.61 (dd, 3H, J_{2-1} = 6.9, J_{2-P} = 0.9, 2), 1.61 ppm (dd, 3H, J_{2-1} = 6.9, J_{2-P} = 0.8, 2). ^{13}C NMR (100 MHz, CDCl$_3$, 25 °C): δ = 169.37 (d, J_{3-P} = 4.9, 3), 169.34 (d, J_{3-P} = 5.2, 3), 153.83 (5''), 150.23–150.43 (m, 1'), 148.26 and 148.25 (4''), 148.16 (d, J_{1-a} = 7.2, 1'''), 148.07 (d, J_{1-a} = 7.6, 1'''), 139.54 and 139.51 (2''), 135.29 (d, J_{4-P} = 1.5, 4'''), 135.29 (d, J_{4-P} = 1.5, 4'''), 130.26 (3'''), 130.15 (3'), 126.47 and 126.45 (1''), 129.83 (d, J_{4-P} = 1.6, 4'), 125.71 (d, J_{4-P} = 1.6, 4'), 120.15 (d, J_{2-P} = 4.6, 2'), 120.10 (d, J_{2-P} = 4.5, 2'), 119.84 (d, J_{2-a} = 5.2, 2''), 119.78 (d, J_{2-P} = 5.1, 2''), 110.02 and 109.99 (6''), 108.10 and 108.08 (3''), 73.38 (d, J_{1-P} = 5.8, 1), 64.28 (4), 56.64 (5''–O–CH_3), 56.39 (4''–O–CH_3), 20.73 and 20.68 (4'''-CH_3), 19.01 (d, J_{2-P} = 5.7, 2), 18.98 ppm (d, J_{2-P} = 5.9, 2). ^{31}P NMR (161 MHz, CDCl$_3$, 25 °C): δ = −12.27 and −12.29. HRMS (ESI+) calculated for $C_{26}H_{27}O_{10}NP$: 532.13671; found: $[M+H]^+$ 532.13678.

4,5-Dimethoxy-2-nitrobenzyl (2S)-2-((ethoxy(4-nitrophenoxy)phosphoryl) oxy)propanoate (3). 4,5-Dimethoxy-2-nitrobenzyl L-lactate **22** (110 mg, 0.39 mmol, 1.00 equiv.) was dissolved in toluene (4 mL), subsequently adding ethyl dichlorophosphate **23** (50 μL, 0.40 mmol, 1.03 equiv.) and TEA (100 μL, 0.72 mmol, 1.85 equiv.) at 25 °C. The mixture was stirred at 25 °C for 12 h and evaporated to dryness *in vacuo*. The solid residue was dissolved in CH_2Cl_2 (4 mL), subsequently adding 4-nitrophenol (56 mg, 0.40 mmol, 1.03 equiv.) and then TEA (100 μL, 0.72 mmol, 1.85 equiv.). The reaction mixture was stirred at 25 °C for

Chemistry–A European Journal

Full Paper
doi.org/10.1002/chem.202101805

Chemistry Europe
European Chemical
Societies Publishing

12 h. The title compound was isolated after evaporating volatiles by normal-phase flash chromatography (n-hexane/ethyl acetate gradient on silica-gel) followed by reversed-phase flash chromatography (water/acetonitrile gradient on C_{18} silica-gel). Yield **3** (34 mg, 18%) of yellowish oils. NOTE: approximately 1:1 mixture of diastereomers. ^{1}H NMR (500 MHz, [D_6]DMSO, 25 °C): $\delta = 8.27$ (m, 2H, 3''), 8.20 (m, 2H, 3''), 7.70 (s, 1H, 3''), 7.67 (s, 1H, 3''), 7.46 (m, 2H, 2'), 7.43 (m, 2H, 2'), 7.21 (s, 1H, 6''), 7.15 (s, 1H, 6''), 5.42–5.55 (m, 4H, 4), 5.18–5.27 (m, 2H, 1), 4.18–4.27 (m, 4H, $-O-CH_2-CH_3$), 3.89 (s, 3H, 5''$-O-CH_3$), 3.87 (s, 3H, 4''$-O-CH_3$), 3.86 (s, 6H, 4''$-O-CH_3$ and 5''$-O-CH_3$), 1.54 (dd, 3H, $J_{2,1} = 7.0$, $J_{2,P} = 0.5$, 2), 1.45 (dd, 3H, $J_{2,1} = 7.0$, $J_{2,P} = 0.6$, 2), 1.25–1.28 ppm (m, 6H, $-O-CH_2-CH_3$). ^{13}C NMR (125 MHz, [D_6]DMSO, 25 °C): $\delta = 169.35$ (d, $J_{1,P} = 4.5$, 3), 169.19 (d, $J_{1,P} = 4.9$, 3), 154.83 (d, $J_{1'-P} = 6.4$, 1'), 153.20 and 153.15 (5''), 148.16 and 148.11 (4''), 144.43 and 144.28 (4'), 139.74 and 139.67 (2''), 125.91 and 125.74 (3'), 125.25 and 125.21 (1''), 121.01 (d, $J_{2'-P} = 5.5$, 2'), 120.93 (d, $J_{2'-P} = 5.5$, 2'), 111.67 and 111.57 (6''), 108.27 and 108.18 (3''), 72.66 (d, $J_{1-P} = 5.6$, 1), 65.49 (d, $J_{CH_2-P} = 6.0$, $-O-CH_2-CH_3$), 65.38 (d, $J_{CH_2-P} = 6.3$, $-O-CH_2-CH_3$), 63.93 and 63.88 (4), 56.31 and 53.25 and 56.13 and 56.08 (4''$-O-CH_3$ and 5''$-O-CH_3$), 18.80 (d, $J_{2-P} = 5.6$, 2), 18.71 (d, $J_{2-P} = 6.3$, 2), 15.70–15.83 ppm (m, $-O-CH_2-CH_3$). ^{31}P NMR (202 MHz, [D_6]DMSO, 25 °C): $\delta = -8.15$ and -8.18 ppm. HRMS (ESI+) calculated for $C_{20}H_{23}O_{12}N_2NaP$: 537.08808; found: [M + Na]$^+$ 537.08765.

4,5-Dimethoxy-2-nitrobenzyl *(2S)-2-((ethoxy(4-fluorophenoxy) phosphoryl)oxy)propanoate* **(4)**. 4,5-Dimethoxy-2-nitrobenzyl L-lactate **22** (110 mg, 0.39 mmol, 1.00 equiv.) was dissolved in toluene (4 mL), subsequently adding ethyl dichlorophosphate **23** (50 μL, 0.40 mmol, 1.03 equiv.) and TEA (100 μL, 0.72 mmol, 1.85 equiv.) at 25 °C. The mixture was stirred at 25 °C for 12 h and evaporated to dryness *in vacuo*. The solid residue was dissolved in CH_2Cl_2 (4 mL), subsequently adding 4-fluorophenol (45 mg, 0.40 mmol, 1.03 equiv.) and then TEA (100 μL, 0.72 mmol, 1.85 equiv.). The reaction mixture was stirred at 25 °C for 12 h. The title compound was isolated after evaporating volatiles by normal-phase flash chromatography (n-hexane/ethyl acetate gradient on silica-gel) followed by reversed-phase flash chromatography (water/acetonitrile gradient on C_{18} silica-gel). Yield **4** (30 mg, 15%) of slightly yellow oils. NOTE: approximately 1:1 mixture of diastereomers. ^{1}H NMR (500 MHz, [D_6]DMSO, 25 °C): $\delta = 7.71$ (s, 1H, 3''), 7.70 (s, 1H, 3''), 7.23–7.24 (m, 2H, 6''), 7.15–7.23 (m, 8H, 2', 3'), 5.51 (d, 1H, $J_{GEM} = 13.9$, 4a), 5.49 (d, 1H, $J_{GEM} = 13.9$, 4a), 5.48 (d, 1H, $J_{GEM} = 13.9$, 4b), 5.46 (d, 1H, $J_{GEM} = 13.9$, 4b), 5.09–5.19 (m, 2H, 1), 4.12–4.21 (m, 4H, $-O-CH_2-CH_3$), 3.89 (s, 3H, 5''$-O-CH_3$ or 4''$-O-CH_3$), 3.87 (s, 3H, 5''$-O-CH_3$ or 4''$-O-CH_3$), 3.86–3.87 (s, 6H, 4''$-O-CH_3$ and/or 5''$-O-CH_3$), 1.52 (d, 3H, $J_{2,1} = 6.9$, 2), 1.45 (dd, 3H, $J_{2,1} = 6.9$, $J_{2,P} = 0.6$, 2), 1.20–1.25 ppm (m, 6H, $-O-CH_2-CH_3$). ^{13}C NMR (125 MHz, [D_6]DMSO, 25 °C): $\delta = 169.53$ (d, $J_{1,P} = 4.6$, 3), 169.36 (d, $J_{1,P} = 5.4$, 3), 159.12 (d, $J_{4'-F} = 241.5$, 4'), 159.08 (d, $J_{4'-F} = 241.5$, 4'), 153.26 and 153.24 (5''), 148.17 and 148.15 (4''), 146.14–146.31 (m, 1'), 139.74 and 139.71 (2''), 125.40 and 125.38 (1''), 121.68–121.94 (m, 2'), 116.53 (d, $J_{3'-F} = 23.7$, 3'), 116.39 (d, $J_{3'-F} = 23.7$, 3'), 111.62 and 111.57 (6''), 108.32 and 108.29 (3''), 72.28 (d, $J_{1-P} = 5.5$, 1), 72.27 (d, $J_{1-P} = 5.5$, 1), 65.00 (d, $J_{CH_2-P} = 6.2$, $-O-CH_2-CH_3$), 64.91 (d, $J_{CH_2-P} = 6.3$, $-O-CH_2-CH_3$), 63.86 and 63.83 (4), 56.34 and 53.29 and 56.17 and 56.15 (4''$-O-CH_3$ and 5''$-O-CH_3$), 18.89 (d, $J_{2-P} = 5.7$, 2), 18.74 (d, $J_{2-P} = 6.2$, 2), 15.74–15.85 ppm (m, $-O-CH_2-CH_3$). ^{31}P NMR (202 MHz, [D_6]DMSO, 25 °C): $\delta = -7.03$ and -7.31 ppm. HRMS (ESI+) calculated for $C_{20}H_{23}O_{10}NFNaP$: 510.09358; found: [M + Na]$^+$ 510.09302.

4,5-Dimethoxy-2-nitrobenzyl *(2S)-2-((ethoxy(4-methylphenoxy) phosphoryl)oxy)propanoate* **(5)**. Ethyl dichlorophosphate **23** (60 μL, 0.47 mmol, 1.00 equiv.) was dissolved in toluene (3 mL), subsequently adding TEA (100 μL, 0.72 mmol, 1.53 equiv.) and phenol (44 mg, 0.47 mmol, 1.0 equiv.) at 25 °C. The mixture was stirred at 25 °C for 12 h and evaporated to dryness *in vacuo*. Solid residue was dissolved in CH_2Cl_2 (3 mL), adding 4,5-dimethoxy-2-nitrobenzyl L-lactate **22**

(110 mg, 0.39 mmol, 0.83 equiv.) and NMI (35 μL; 0.43 mmol, 0.91 equiv.) at 25 °C. The mixture was stirred at 25 °C for 12 h. The title compound was isolated (direct injection) by normal-phase flash chromatography (n-hexane/ethyl acetate gradient on silica-gel) followed by reversed-phase flash chromatography (water/methanol gradient on C_{18} silica-gel). NOTE: approximately 1:1 mixture of diastereomers. Yield **5** (20 mg, 11%) of a yellowish oil. ^{1}H NMR (500 MHz, [D_6]DMSO, 25 °C): $\delta = 7.72$ (s, 1H, 3''), 7.71 (s, 1H, 3''), 7.40 (m, 2H, 3'), 7.35 (m, 2H, 3'), 7.23 (s, 1H, 6''), 7.17–7.22 (m, 7H, 2', 4', 6''), 5.43–5.54 (m, 4H, 4), 5.08–5.18 (m, 2H, 1), 4.12–4.20 (m, 4H, $-O-CH_2-CH_3$), 3.89 (s, 3H, 4''$-O-CH_3$ or 5''$-O-CH_3$), 3.88 (s, 3H, 4''$-O-CH_3$ or 5''$-O-CH_3$), 3.87 (s, 3H, 4''$-O-CH_3$ or 5''$-O-CH_3$), 3.87 (s, 3H, 4''$-O-CH_3$ or 5''$-O-CH_3$), 1.52 (dd, 3H, $J_{2,1} = 6.8$, $J_{2,P} = 0.4$, 2), 1.45 (dd, 3H, $J_{2,1} = 6.9$, $J_{2,P} = 0.6$, 2), 1.23 (td, 3H, $J_{CH_3-CH_2} = 7.1$, $J_{CH_3-P} = 1.0$, $-O-CH_2-CH_3$), 1.22 ppm (td, 3H, $J_{CH_3-CH_2} = 7.2$, $J_{CH_3-P} = 1.0$, $-O-CH_2-CH_3$). ^{13}C NMR (125 MHz, [D_6]DMSO, 25 °C): $\delta = 169.53$ (d, $J_{1,P} = 4.6$, 3), 169.35 (d, $J_{1,P} = 5.3$, 3), 153.25 and 153.23 (5''), 150.12 (d, $J_{1'-P} = 6.4$, 1'), 148.13 and 148.11 (4''), 139.70 and 139.68 (2''), 129.96 and 129.84 (3'), 125.42 (1''), 125.35 and 125.22 (4'), 119.99 (d, $J_{2'-P} = 5.1$, 2'), 119.89 (d, $J_{2'-P} = 4.5$, 2'), 111.56 and 111.51 (6''), 108.30 and 108.28 (3''), 72.17 (d, $J_{1-P} = 5.5$, 1), 72.14 (d, $J_{1-P} = 5.5$, 1), 64.83 (d, $J_{CH_2-P} = 6.4$, $-O-CH_2-CH_3$), 64.76 (d, $J_{CH_2-P} = 6.1$, $-O-CH_2-CH_3$), 63.80 and 63.76 (4), 56.31 (4''$-O-CH_3$ or 5''$-O-CH_3$), 56.26 (4''$-O-CH_3$ or 5''$-O-CH_3$), 56.11–56.16 (m, 4''$-O-CH_3$ or/and 5''$-O-CH_3$), 18.87 (d, $J_{2-P} = 5.5$, 2), 18.71 (d, $J_{2-P} = 6.3$, 2), 15.70–15.84 ppm (m, $-O-CH_2-CH_3$). ^{31}P NMR (202 MHz, [D_6]DMSO, 25 °C): $\delta = -7.14$ and -7.47 ppm. HRMS (ESI+) calculated for $C_{20}H_{24}O_{10}NNaP$: 492.10300; found: [M + Na]$^+$ 492.10263.

4,5-Dimethoxy-2-nitrobenzyl *(2S)-2-((ethoxy(4-methylphenoxy) phosphoryl)oxy)propanoate* **(6)**. 4,5-Dimethoxy-2-nitrobenzyl L-lactate **22** (110 mg, 0.39 mmol, 1.00 equiv.) was dissolved in toluene (3 mL), subsequently adding ethyl dichlorophosphate **23** (50 μL, 0.40 mmol, 1.03 equiv.) and TEA (100 μL, 0.72 mmol, 1.85 equiv.) at 25 °C. The mixture was stirred at 25 °C for 12 h and evaporated to dryness *in vacuo*. The solid residue was dissolved in CH_2Cl_2 (3 mL), adding 4-methylphenol (45 mg, 0.42 mmol, 1.08 equiv.) and then TEA (100 μL, 0.72 mmol, 1.85 equiv.). The reaction mixture was stirred at 25 °C for 12 h. The title compound was isolated after evaporating volatiles by normal-phase flash chromatography (n-hexane/ethyl acetate gradient on silica-gel) followed by reversed-phase flash chromatography (water/acetonitrile gradient on C_{18} silica-gel). Yield **6** (21 mg, 11%) of slightly yellow oil. NOTE: approximately 1:1 mixture of diastereomers. ^{1}H NMR (500 MHz, [D_6]DMSO, 25 °C): $\delta = 7.72$ (s, 1H, 3''), 7.71 (s, 1H, 3''), 7.04–7.23 (m, 10H, 2', 3', 6''), 5.52 (d, 1H, $J_{GEM} = 13.8$, 4a), 5.47 (d, 1H, $J_{GEM} = 13.8$, 4b), 5.50 (d, 1H, $J_{GEM} = 14.0$, 4a), 5.45 (d, 1H, $J_{GEM} = 14.0$, 4b), 5.06–5.16 (m, 2H, 1), 4.11–4.19 (m, 4H, $-O-CH_2-CH_3$), 3.89 (s, 3H, 5''$-O-CH_3$), 3.87 (s, 3H, 4''$-O-CH_3$), 3.86 (s, 3H, 5''$-O-CH_3$), 2.27 (s, 3H, 4'-CH_3), 2.24 (s, 3H, 4'-CH_3), 1.51 (d, 3H, $J_{2,1} = 6.9$, 2), 1.45 (dd, 3H, $J_{2,1} = 6.8$, $J_{2,P} = 0.5$, 2), 1.22 (td, 3H, $J_{CH_3-CH_2} = 7.0$, $J_{CH_3-P} = 1.0$, $-O-CH_2-CH_3$), 1.21 ppm (td, 3H, $J_{CH_3-CH_2} = 7.1$, $J_{CH_3-P} = 1.0$, $-O-CH_2-CH_3$). ^{13}C NMR (125 MHz, [D_6]DMSO, 25 °C): $\delta = 169.56$ (d, $J_{1,P} = 4.8$, 3), 169.36 (d, $J_{1,P} = 5.3$, 3), 153.25 (5''), 148.13 and 148.10 (4''), 148.00–147.86 (m, 1'), 139.69 and 139.65 (2''), 134.51 (d, $J_{4'-P} = 1.5$, 4'), 134.37 (4'), 130.22 and 130.09 (3'), 125.45 and 125.43 (1''), 119.73 (d, $J_{2'-P} = 4.6$, 2'), 119.64 (d, $J_{2'-P} = 4.6$, 2'), 111.55 and 111.47 (6''), 108.29 and 108.26 (3''), 72.12 (d, $J_{1-P} = 5.5$, 1), 72.10 (d, $J_{1-P} = 5.5$, 1), 64.75 (d, $J_{CH_2-P} = 6.2$, $-O-CH_2-CH_3$), 64.68 (d, $J_{CH_2-P} = 6.2$, $-O-CH_2-CH_3$), 63.79 and 63.75 (4), 56.31 and 53.26 (4''$-O-CH_3$ or 5''$-O-CH_3$), 56.12–56.16 (m, 4''$-O-CH_3$ or/and 5''$-O-CH_3$), 20.26 and 20.21 (4'-CH_3), 18.88 (d, $J_{2-P} = 5.5$, 2), 18.74 (d, $J_{2-P} = 6.3$, 2), 15.81 (d, $J_{CH_2-P} = 6.3$, $-O-CH_2-CH_3$), 15.76 ppm (d, $J_{CH_2-P} = 6.3$, $-O-CH_2-CH_3$). ^{31}P NMR (202 MHz, [D_6]DMSO, 25 °C): $\delta = -6.97$ and -7.26 ppm. HRMS (ESI+) calculated for $C_{21}H_{26}O_{10}NNaP$: 506.11865; found: [M + Na]$^+$ 506.11847.

4,5-Dimethoxy-2-nitrobenzyl *(2S)-2-((ethoxy(4-methoxyphenoxy) phosphoryl)oxy)propanoate* **(7)**. 4,5-Dimethoxy-2-nitrobenzyl L-lactate **22** (500 mg, 1.75 mmol, 1.00 equiv.) was dissolved in toluene (10 mL),

Chem. Eur. J. 2021, 27, 12763–12775   www.chemeurj.org

12770

© 2021 Wiley-VCH GmbH

subsequently adding ethyl dichlorophosphate **23** (250 μL, 1.98 mmol, 1.13 equiv.) and TEA (300 μL, 2.16 mmol, 1.23 equiv.) at 25 °C. The mixture was stirred at 25 °C for 12 h and evaporated to dryness *in vacuo*. The solid residue was dissolved in CH_2Cl_2 (10 mL), adding 4-methoxyphenol (90 mg, 0.73 mmol, 0.42 equiv.) and then TEA (100 μL, 0.72 mmol, 0.41 equiv.). The reaction mixture was stirred at 25 °C for 12 h. The title compound was isolated after evaporating volatiles by normal-phase flash chromatography (CH_2Cl_2/methanol gradient on silica-gel) followed by reversed-phase flash chromatography (water/acetonitrile gradient on C_{18} silica-gel). Yield **7** (56 mg, 16 %) of a yellow oil. NOTE: approximately 1:0.7 mixture of diastereomers. ^1H NMR (400 MHz, $CDCl_3$, 25 °C): $\delta = 7.74$ and 7.73 (s, 2H, 3''), 7.10–7.15 (m, 4H, 2'), 7.08 (s, 1H, 6''), 7.05 (s, 1H, 6''), 6.84 (m, 2H, 3'), 6.77 (m, 2H, 3'), 5.65 (d, 1H, $J_{GEM} = 15.1$, 4a), 5.60 (d, 1H, $J_{GEM} = 15.1$, 4a), 5.63 (d, 1H, $J_{GEM} = 14.9$, 4a), 5.55 (d, 1H, $J_{GEM} = 14.9$, 4b), 5.03–5.16 (m, 2H, 1), 4.19–4.31 (m, 4H, $-O-CH_2-CH_3$), 4.00 and 3.97 and 3.97 and 3.95 (s, 12H, 5''-O-CH_3 and 4''-O-CH_3), 3.79 (s, 3H, 4'-O-CH_3), 3.75 (s, 3H, 4'-O-CH_3), 1.67 (dd, 3H, $J_{2-1} = 6.9$, $J_{2-P} = 0.6$, 2), 1.56 (dd, 3H, $J_{2-1} = 7.0$, $J_{2-P} = 0.9$, 2), 1.35 (td, 3H, $J_{CH_3-CH_2} = 6.3$, $J_{CH_3-P} = 1.2$, $-O-CH_2-CH_3$), 1.33 ppm (td, 3H, $J_{CH_3-CH_2} = 6.3$, $J_{CH_3-P} = 1.2$, $-O-CH_2-CH_3$). ^{13}C NMR (100 MHz, $CDCl_3$, 25 °C): $\delta = 169.90$ (d, $J_{1-P} = 4.2$, 3), 169.53 (d, $J_{1-P} = 5.4$, 3), 156.77 and 156.75 (4'), 153.79 and 153.76 (5'' or 4''), 148.29 and 148.22 (4'' or 5''), 143.97 (d, $J_{1-P} = 7.6$, 1'), 143.92 (d, $J_{1-P} = 7.2$, 1'), 139.61 and 139.51 (2''), 126.44 (1''), 120.96 (d, $J_{2-P} = 4.6$, 2'), 120.88 (d, $J_{2-P} = 4.6$, 2'), 114.58 and 114.47 (3'), 110.11 and 110.03 (6''), 108.14 and 108.07 (3''), 72.50 (d, $J_{1-P} = 5.3$, 1), 72.48 (d, $J_{1-P} = 5.1$, 1), 65.18 (d, $J_{CH_2-P} = 6.2$, $-O-CH_2-CH_3$), 64.94 (d, $J_{CH_2-P} = 6.1$, $-O-CH_2-CH_3$), 64.18 and 64.14 (4), 56.67 and 56.60 and 56.35 (5''-O-CH_3 and 4''-O-CH_3), 55.56 and 55.50 (4'-O-CH_3), 19.16 (d, $J_{2-P} = 5.4$, 2), 18.99 (d, $J_{2-P} = 6.2$, 2), 16.00 (d, $J_{CH_3-P} = 6.4$, $-O-CH_2-CH_3$), 15.95 ppm (d, $J_{CH_3-P} = 6.6$, $-O-CH_2-CH_3$). ^{31}P NMR (161 MHz, $CDCl_3$, 25 °C): $\delta = -M \rightarrow 6.48$, -6.67 ppm. HRMS (ESI+) calculated for $C_{21}H_{27}O_{11}NP$: 500.13162; found: $[M+H]^+$ 500.13106.

4,5-dimethoxy-2-nitrobenzyl (2S)-2-((ethoxy(phenoxy)phosphoryl)oxy)-3-methylbutanoate (8). Ethyl dichlorophosphate **23** (50 μL, 0.40 mmol, 1.00 equiv.) was dissolved in toluene (3 mL), subsequently adding TEA (100 μL, 0.72 mmol, 1.80 equiv.) and phenol (40 mg, 0.43 mmol, 1.08 equiv.) at 25 °C. The mixture was stirred at 25 °C for 12 h and evaporated to dryness *in vacuo*. The solid residue was dissolved in CH_2Cl_2 (2 mL), adding 4,5-dimethoxy-2-nitrobenzyl (S)-2-hydroxy-3-methylbutanoate **28** (120 mg, 0.38 mmol, 0.95 equiv.) and DMAP (100 mg, 0.82 mmol, 2.05 equiv.) at 25 °C. The mixture was stirred at 25 °C for 12 h. The title compound was isolated (direct injection) by normal-phase flash chromatography (n-hexane/ethyl acetate gradient on silica-gel) followed by reversed-phase flash chromatography (water/methanol gradient on C_{18} silica-gel). Yield **8** (20 mg, 11 %) of a yellowish oil. NOTE: approximately 1:1 mixture of diastereomers. ^1H NMR (400 MHz, $CDCl_3$, 25 °C): $\delta = 7.74$ (s, 1H, 3''), 7.72 (s, 1H, 3''), 7.25–7.31 (m, 4H, 3'), 7.17–7.25 (m, 4H, 2'), 7.13–7.17 (m, 2H, 4'), 7.12 (s, 1H, 6''), 7.07 (s, 1H, 6''), 5.66 (d, $J_{GEM} = 14.8$, 3a), 5.60 (d, $J_{GEM} = 14.8$, 3b), 5.60 (d, $J_{GEM} = 15.0$, 3a), 5.54 (d, $J_{GEM} = 15.0$, 3b), 4.81–4.88 (m, 2H, 1), 4.21–4.33 (m, 4H, $-O-CH_2-CH_3$), 4.02 (s, 3H, 5''-OCH_3), 3.95 (s, 3H, 5''-OCH_3), 3.98 (s, 3H, 4''-OCH_3), 3.97 (s, 3H, 4''-OCH_3), 2.34 (m, 1H, CH^{iPr}), 2.27 (m, 1H, CH^{iPr}), 1.35 (td, 3H, $J_{CH_3-CH_2} = 7.1$, $J_{CH_3-P} = 1.1$, $-O-CH_2-CH_3$), 1.34 (td, 3H, $J_{CH_3-CH_2} = 7.1$, $J_{CH_3-P} = 1.1$, $-O-CH_2-CH_3$), 1.09 (d, 3H, $J_{CH_3(iPr)-CH(iPr)} = 6.9$, CH_3^{iPr}), 1.03 (d, 3H, $J_{CH_3(iPr)-CH(iPr)} = 6.9$, CH_3^{iPr}), 0.97 ppm (d, 6H, $J_{CH_3(iPr)-CH(iPr)} = 6.8$, CH_3^{iPr}). ^{13}C NMR (100 MHz, $CDCl_3$, 25 °C): $\delta = 169.23$ (d, $J_{2-P} = 1.6$, 2), 168.86 (d, $J_{2-P} = 2.1$, 2), 153.83 and 153.78 (5''), 148.31 and 148.23 (4''), 150.60 (1'), 139.67 and 139.60 (2''), 129.68 and 129.56 (3'), 126.57 and 126.53 (1''), 125.17 (d, $J_{4'-P} = 1.6$, 4'), 125.11 (d, $J_{4'-P} = 1.5$, 4'), 120.02 (d, $J_{2'-P} = 4.6$, 2'), 119.97 (d, $J_{2'-P} = 4.7$, 2'), 110.37 and 110.27 (6''), 108.12 and 108.04 (3''), 80.78 (d, $J_{1-P} = 6.5$, 1), 65.24 (d, $J_{CH_2-P} = 6.3$, $-O-CH_2-CH_3$), 64.99 (d, $J_{CH_2-P} = 6.3$, $-O-CH_2-CH_3$), 64.01 and 63.93 (3), 56.74 and 56.66 (5''-O-CH_3), 56.39 and 56.38 (4''-O-CH_3), 31.78 (d, $J_{CH(iPr)-P}$

6.9, CH^{iPr}), 31.69 (d, $J_{CH(iPr)-P} = 7.1$, CH^{iPr}), 18.47 and 18.37 and 16.76 and 16.61 (CH_3^{iPr}), 16.03 (d, $J_{CH_3-P} = 7.4$, $-O-CH_2-CH_3$), 15.94 ppm (d, $J_{CH_3-P} = 6.9$, $-O-CH_2-CH_3$). ^{31}P NMR (161 MHz, $CDCl_3$, 25 °C): $\delta = -6.67$ ppm. HRMS (ESI+) calculated for $C_{22}H_{28}O_{10}NNaP$: 520.13430; found: $[M+Na]^+$ 520.13434.

4,5-Dimethoxy-2-nitrobenzyl 2-((ethoxy(phenoxy)phosphoryl)oxy)-2-methylpropanoate (9). Ethyl dichlorophosphate **23** (60 μL, 0.47 mmol, 1.00 equiv.) was dissolved in toluene (3 mL), subsequently adding TEA (100 μL, 0.72 mmol, 1.80 equiv.) and phenol (45 mg, 0.48 mmol, 1.02 equiv.) at 25 °C. The mixture was stirred at 25 °C for 12 h and evaporated to dryness *in vacuo*. The solid residue was dissolved in CH_2Cl_2 (3 mL), adding 4,5-dimethoxy-2-nitrobenzyl (2-methyl)propanoate **29** (140 mg, 0.47 mmol, 1.00 equiv.) and NMI (50 μL, 0.61 mmol, 1.30 equiv.) at 25 °C. The mixture was stirred at 25 °C for 12 h and evaporated to dryness *in vacuo*. The title compound was isolated by normal-phase flash chromatography (n-hexane/ethyl acetate gradient on silica-gel) followed by reversed-phase flash chromatography (water/methanol gradient on C_{18} silica-gel). Yield **9** (14 mg, 6 %) of a yellowish oil. ^1H NMR (400 MHz, $CDCl_3$, 25 °C): $\delta = 7.72$ (s, 1H, 3''), 7.30 (m, 2H, 3'), 7.20 (m, 2H, 2'), 7.15 (m, 1H, 4'), 7.14 (s, 1H, 6''), 5.63 (dd, $J_{GEM} = 14.9$, $J_{3a-C} = 0.6$, 3a), 5.57 (dd, $J_{GEM} = 14.9$, $J_{3b-C} = 0.6$, 3b), 4.23 (m, 2H, $-O-CH_2-CH_3$), 3.97 (s, 3H, 5''-O-CH_3), 3.97 (s, 3H, 4''-O-CH_3), 1.76 (s, 3H, 1-CH_3), and 1.75 (s, 3H, 1-CH_3), 1.32 ppm (td, $J_{CH_3-CH_2} = 7.1$, $J_{CH_3-P} = 1.2$, $-O-CH_2-CH_3$). ^{13}C NMR (100 MHz, $CDCl_3$, 25 °C): $\delta = 171.64$ (d, $J_{2-P} = 3.9$, 2), 153.75 (5''), 150.61 (d, $J_{1-P} = 6.9$, 1'), 148.20 (4''), 139.62 (2''), 129.54 (3'), 126.69 (1''), 124.99 (4'), 120.10 (d, $J_{2'-P} = 4.7$, 2'), 110.40 (6''), 108.04 (3''), 82.05 (d, $J_{1-P} = 6.2$, 1), 64.71 (d, $J_{CH_2-P} = 6.2$, $-O-CH_2-CH_3$), 64.50 (3), 56.69 and 56.36 (5''-O-CH_3 and 4''-O-CH_3), 26.56 (d, $J_{CH_3-P} = 5.5$, 1-CH_3), 26.08 (d, $J_{CH_3-P} = 4.1$, 1-CH_3), 15.92 ppm (d, $J_{CH_3-P} = 7.0$, $-O-CH_2-CH_3$). ^{31}P NMR (161 MHz, $CDCl_3$, 25 °C): $\delta = -10.38$ ppm. HRMS (ESI+) calculated for $C_{21}H_{27}O_{10}NP$: 484.13671; found: $[M+H]^+$ 484.13656.

4,5-Dimethoxy-2-nitrobenzyl (2S)-2-(((4-fluorophenoxy)(phenoxy)phosphoryl)oxy)-3-methylbutanoate (10). Phenyl dichlorophosphate **17** (60 μL, 0.40 mmol, 1.00 equiv.) was dissolved in toluene (3 mL), subsequently adding TEA (100 μL, 0.72 mmol, 1.80 equiv.) and 4-fluorophenol (45 mg, 0.40 mmol, 1.00 equiv.) at 25 °C. The mixture was stirred at 25 °C for 12 h and evaporated to dryness *in vacuo*. The solid residue was dissolved in CH_2Cl_2 (2 mL), adding 4,5-dimethoxy-2-nitrobenzyl (S)-2-hydroxy-3-methylbutanoate **28** (120 mg, 0.38 mmol, 0.95 equiv.) and DMAP (100 mg, 0.82 mmol, 2.05 equiv.) at 25 °C. The mixture was stirred at 25 °C for 12 h. The title compound was isolated (direct injection) by normal-phase flash chromatography (hexanes/ethyl acetate gradient on silica-gel) followed by reversed-phase flash chromatography (water/methanol gradient on C_{18} silica-gel). Yield **10** (149 mg, 70 %) of a yellowish oil. NOTE: approximately 1:1 mixture of diastereomers. ^1H NMR (400 MHz, $CDCl_3$, 25 °C): $\delta = 7.68$ (s, 1H, 3''), 7.68 (s, 1H, 3''), 7.32 (m, 2H, 3'), 7.26 (m, 2H, 3'), 7.10–7.22 (m, 10H, 2', 4', 2''), 6.90–7.03 (m, 6H, 6'', 3'''), 5.51–5.61 (m, 4H, 3), 4.92 (dd, $J_{1-CH(iPr)} = 4.2$, $J_{1-P} = 7.6$, 1), 4.91 (dd, $J_{1-CH(iPr)} = 4.2$, $J_{1-P} = 7.4$, 1), 3.92–3.94 (m, 9H, 4''-O-CH_3), 3.91 (s, 3H, 5''-O-CH_3), 2.30 (m, 2H, CH^{iPr}), 0.94–1.00 ppm (m, 12H, CH_3^{iPr}). ^{13}C NMR (100 MHz, $CDCl_3$, 25 °C): $\delta = 168.56$ (d, $J_{2-P} = 1.9$, 2), 168.46 (d, $J_{2-P} = 1.9$, 2), 159.70 (dd, $J_{4'-F} = 245.0$, $J_{4'-P} = 1.2$, 4'''), 159.66 (dd, $J_{4'-F} = 244.3$, $J_{4'-P} = 1.5$, 4'''), 153.58 (5''), 150.22 (d, $J_{1'-P} = 6.7$, 1'), 150.15 (d, $J_{1'-P} = 6.6$, 1'), 148.22 and 148.18 (4''), 146.04–146.23 (m, 1'''), 139.62 and 139.58 (2''), 129.71 and 129.54 (3'), 126.09 and 126.00 (1''), 125.47 and 125.39 (4'), 121.56 (dd, $J_{2'-F} = 8.6$, $J_{2'-P} = 4.8$, 2'''), 121.35 (dd, $J_{2'-F} = 8.4$, $J_{2'-P} = 5.4$, 2'''), 119.97 (d, $J_{2'-P} = 4.7$, 2'), 119.82 (d, $J_{2'-P} = 4.6$, 2'), 116.25 (d, $J_{3'-F} = 23.7$, 3'''), 116.03 (d, $J_{3'-F} = 23.8$, 3'''), 110.35 and 110.30 (6''), 107.94 (3''), 81.59 (d, $J_{1-P} = 6.6$, 1), 64.01 (3), 56.49 (5''-O-CH_3), 56.22 (4''-O-CH_3), 31.59 (d, $J_{CH(iPr)-P} = 7.2$, CH^{iPr}), 18.23 and 18.20 and 16.48 and 16.44 ppm (CH_3^{iPr}). ^{31}P NMR (161 MHz, $CDCl_3$, 25 °C): $\delta = -11.96$ and

−11.98 ppm. HRMS (ESI+) calculated for $C_{26}H_{28}O_{10}NFP$: 564.14294; found: $[M+H]^+$ 564.14368.

4,5-Dimethoxy-2-nitrobenzyl (2S)-3-methyl-2-((phenoxy(p-tolyloxy)phosphoryl)oxy)butanoate (11). Phenyl dichlorophosphate **17** (66 µL, 0.44 mmol, 1.00 equiv.) was dissolved in toluene (3 mL), subsequently adding TEA (66 µL, 0.47 mmol, 1.07 equiv.) and 4-methylphenol (45 mg, 0.42 mmol, 0.95 equiv.) at 25 °C. The mixture was stirred at 25 °C for 12 h and evaporated to dryness *in vacuo*. The solid residue was dissolved in CH_2Cl_2 (2 mL), adding 4,5-dimethoxy-2-nitrobenzyl (S)-2-hydroxy-3-methylbutanoate **28** (100 mg, 0.32 mmol, 0.73 equiv.) and DMAP (60 mg, 0.49 mmol, 1.11 equiv.) at 25 °C. The mixture was stirred at 25 °C for 12 h. The title compound was isolated (direct injection) by normal-phase flash chromatography (hexanes/ethyl acetate gradient on silica-gel) followed by reversed-phase flash chromatography (water/methanol gradient on C_{18} silica-gel). Yield **11** (103 mg, 58%) of yellowish oil. NOTE: approximately 1:1 mixture of diastereomers. ^1H NMR (400 MHz, $CDCl_3$, 25 °C): δ = 7.70 (s, 1H, 3"), 7.70 (s, 1H, 3"), 7.25–7.35 (m, 4H, 3'), 7.18–7.24 (m, 4H, 2'), 7.11–7.16(m, 2H, 4'), 7.02–7.11 (m, 10H, 2"', 3"', 6"), 5.55–5.57 (m, 4H, 3), 4.92–4.95 (m, 2H, 1), 3.95 (s, 6H, 4"'-O-CH_3), 3.92 (s, 3H, 5"'-O-CH_3), 3.92 (s, 3H, 5"'-O-CH_3), 2.25–2.36 (m, 8H, CH^{iPr}, 4"''-CH_3) 0.97–1.01 ppm (m, 12H, CH_3^{iPr}). ^{13}C NMR (100 MHz, $CDCl_3$, 25 °C): δ = 168.64 (d, J_{2-P} = 1.9, 2), 153.72 (5"), 150.40 (d, J_{1-P} = 7.6, 1'), 150.36 (d, J_{1-P} = 7.0, 1'), 148.09–148.23 (m, 1"', 4"), 139.49 and 139.45 (2"), 135.06 (4"''), 135.00 (d, J_{4-P} = 1.2, 4"'), 130.12 and 129.97 (3"'), 129.70 and 129.53 (3'), 126.46 and 126.43 (1"), 125.37 and 125.30 (4'), 120.08 (d, J_{2-P} = 4.7, 2'), 119.92 (d, J_{2-P} = 4.9, 2'), 119.78 (d, J_{2-P} = 4.7, 2"'), 119.59 (d, J_{2-P} = 4.7, 2"'), 110.17 and 110.13 (6"), 107.93 and 107.92(3"), 81.53 (d, J_{1-P} = 6.9, 1), 63.97 (3), 56.59 (5"'-O-CH_3), 56.28 (4"'-O-CH_3), 31.70 (d, $J_{CH(iPr)-P}$ = 7.1, CH^{iPr}), 31.69 (d, $J_{CH(iPr)-P}$ = 7.1, CH^{iPr}), 20.62 and 20.58 (4"''-CH_3), 18.28 and 18.26 and 16.61 and 16.59 ppm (CH_3^{iPr}). ^{31}P NMR (161 MHz, $CDCl_3$, 25 °C): δ = −11.91 and −11.96. HRMS (ESI+) calculated for $C_{27}H_{31}O_{10}NP$: 560.16801; found: $[M+H]^+$ 560.16748.

4,5-Dimethoxy-2-nitrobenzyl 2-(((4-fluorophenoxy)(phenoxy)phosphoryl)oxy)-2-methylpropanoate (12). Phenyl dichlorophosphate **17** (60 µL, 0.40 mmol, 1.00 equiv.) was dissolved in toluene (3 mL), subsequently adding TEA (100 µL, 0.72 mmol, 1.80 equiv.) and 4-fluorophenol (45 mg, 40 mmol, 1.00 equiv.) at 25 °C. The mixture was stirred at 25 °C for 12 h and evaporated to dryness *in vacuo*. The solid residue was dissolved in CH_2Cl_2 (3 mL), adding 4,5-dimethoxy-2-nitrobenzyl (2-methyl)propanoate **29** (120 mg, 0.40 mmol, 1.00 equiv.), TEA (100 µL, 0.72 mmol, 1.80 equiv.) and a catalytic amount of DMAP at 25 °C. The mixture was stirred at 25 °C for 1 h and evaporated to dryness *in vacuo*. The title compound was isolated by normal-phase flash chromatography (n-hexane/ethyl acetate gradient on silica-gel) followed by reversed-phase flash chromatography (water/methanol gradient on C_{18} silica-gel). Yield **12** (21 mg, 10%) of a yellowish oil. ^1H NMR (400 MHz, $CDCl_3$, 25 °C): δ = 7.70 (s, 1H, 3"), 7.31 (m, 2H, 3'), 7.14–7.21 (m, 5H, 2', 4', 2"'), 7.07 (s, 1H, 6"), 6.98 (m, 2H, 3"'), 5.58 (dd, J_{GEM} = 14.9, J_{3a-C} = 0.6, 3a), 5.54 (dd, J_{GEM} = 14.9, J_{3b-C} = 0.6, 3b), 3.96 (s, 3H, 4"'-O-CH_3), 3.94 (s, 3H, 5"'-O-CH_3), 1.76 (s, 3H, 1-CH_3), 1.75 ppm (s, 3H, 1-CH_3). ^{13}C NMR (100 MHz, $CDCl_3$, 25 °C): δ = 171.23 (d, J_{2-P} = 3.9, 2), 159.77 (dd, J_{4-F} = 244.0, J_{4-P} = 1.5, 4"'), 153.62 (5"), 150.36 (d, J_{1-P} = 7.4, 1'), 148.25 (4"), 146.30 (dd, J_{1-P} = 7.7, J_{1-F} = 3.1, 1"'), 139.71 (2"), 129.64 (3'), 126.26 (1"), 125.38 (d, J_{4-P} = 1.4, 4'), 121.61 (dd, J_{2-F} = 8.5, J_{2-P} = 4.7, 2"'), 120.09 (d, J_{2-P} = 4.7, 2'), 116.18 (d, J_{3-F} = 23.9, 3"'), 110.54 (6"), 108.02 (3"), 83.28 (d, J_{1-P} = 6.6, 1), 64.67 (3), 56.59 (5"'-O-CH_3), 56.32 (4"'-O-CH_3), 26.39 (d, J_{CH_3-P} = 5.0, 1-CH_3), 26.28 ppm (d, J_{CH_3-P} = 4.9, 1-CH_3). ^{31}P NMR (161 MHz, $CDCl_3$, 25 °C): δ = −15.92 and −15.93 ppm. HRMS (ESI+) calculated for $C_{25}H_{25}O_{10}NFP$: 550.12729; found: $[M+H]^+$ 550.12715.

4,5-Dimethoxy-2-nitrobenzyl 2-(((phenoxy)(p-tolyloxy)phosphoryl)oxy)-2-methylpropanoate (13). Phenyl dichlorophosphate **17** (66 µL,

0.44 mmol, 1.00 equiv.) was dissolved in toluene (3 mL), subsequently adding TEA (66 µL, 0.47 mmol, 1.07 equiv.) and 4-methylphenol (45 mg, 0.42 mmol, 0.95 equiv.) at 25 °C. The mixture was stirred at 25 °C for 12 h and evaporated to dryness *in vacuo*. The solid residue was dissolved in CH_2Cl_2 (2 mL), adding 4,5-dimethoxy-2-nitrobenzyl (2-methyl)propanoate **29** (100 mg, 0.33 mmol, 0.75 equiv.) and DMAP (60 mg, 0.49 mmol, 1.11 equiv.) at 25 °C. The mixture was stirred at 25 °C for 12 h and evaporated to dryness *in vacuo*. The title compound was isolated by normal-phase flash chromatography (n-hexane/ethyl acetate gradient on silica-gel) followed by reversed-phase flash chromatography (water/methanol gradient on C_{18} silica-gel). Yield 13 (97 mg, 54%) of a yellowish oil. ^1H NMR (400 MHz, $CDCl_3$, 25 °C): δ = 7.70 (s, 1H, 3"), 7.30 (m, 2H, 3'), 7.20 (m, 2H, 2'), 7.16 (m, 4'), 7.11 (s, 1H, 6"), 7.07–7.08 (m, 4H, 2"', 3"'), 5.56 (s, 2H, 3), 3.96 (s, 3H, 4"'-O-CH_3), 3.92 (s, 3H, 5"'-O-CH_3), 2.30 (d, 3H, J_{CH_3-9} = 0.8, 4"''-CH_3), 1.76 ppm (s, 6H, 1-(CH_3)$_2$). ^{13}C NMR (100 MHz, $CDCl_3$, 25 °C): δ = 171.34 (d, J_{2-P} = 4.2, 2), 153.71 (5"), 150.48 (d, J_{1-P} = 7.7, 1'), 148.25 (d, J_{1-P} = 7.7, 1"'), 148.12 (4"), 139.51 (2"), 134.89 (4"'), 130.00 (3"'), 129.56 (3'), 126.61 (1"), 125.22 (d, J_{4-P} = 1.0, 4'), 120.14 (d, J_{2-P} = 4.7, 2'), 119.81 (d, J_{2-P} = 4.7, 2"'), 110.31 (6"), 107.93 (3"), 83.06 (d, J_{1-P} = 6.6, 1), 64.57 (3), 56.62 (5"'-O-CH_3), 56.30 (4"'-O-CH_3), 26.33 (d, J_{CH_3-P} = 5.2, 1-(CH_3)$_2$), 20.63 ppm (4"''-CH_3). ^{31}P NMR (161 MHz, $CDCl_3$, 25 °C): δ = −15.90 ppm. HRMS (ESI+) calculated for $C_{26}H_{28}O_{10}NP$: 546.15236; found: $[M+H]^+$ 546.15200.

4,5-Dimethoxy-2-nitrobenzyl (S)-2-((diethoxyphosphoryl)oxy)propanoate (14). 4,5-Dimethoxy-2-nitrobenzyl ʟ-lactate **22** (114 mg, 0.40 mmol, 1.00 equiv.) was dissolved in dry CH_2Cl_2 (3 mL), subsequently adding diethyl chlorophosphate **30** (60 µL, 0.41 mmol, 1.04 equiv.) and TEA (100 µL, 0.72 mmol, 1.80 equiv.) at 25 °C. The mixture was stirred at 25 °C for 36 h and directly injected to normal-phase flash chromatography system (hexanes/ethyl acetate gradient on silica-gel), followed by reversed-phase flash chromatography (water/methanol gradient on C_{18} silica-gel). Yield 14 (18 mg, 11%) of a yellowish oil. ^1H NMR (400 MHz, $CDCl_3$, 25 °C): δ = 7.75 (s, 1H, 3"), 7.11 (s, 1H, 6"), 5.66 (d, J_{GEM} = 14.8, 3a), 5.62 (d, J_{GEM} = 14.8, 3b), 5.03 (m, 1), 4.10–4.22 (m, 4H, O-CH_2-CH_3), 4.03 (s, 3H, 5"'-O-CH_3), 3.98 (s, 3H, 4"'-O-CH_3), 1.64 (dd, 3H, J_{2-1} = 6.9, J_{2-P} = 0.6, 1-Me), 1.36 (td, 3H, J_{7-T} = 7.1, J_{CH_3-P} = 1.0, O-CH_2-CH_3), 1.33 ppm (td, 3H, J_{CH_3-T} = 7.1, J_{7-P} = 1.0, O-CH_2-CH_3). ^{13}C NMR (100 MHz, $CDCl_3$, 25 °C): δ = 170.13 (d, J_{2-P} = 4.8, 2), 153.81 (5"), 148.32 (4"), 139.6 (2"), 126.53 (1"), 110.16 (6"), 108.19 (3"), 71.72 (d, J_{1-P} = 5.4, 1), 64.32 (d, J_{CH_2-P} = 5.9, O-CH_2-CH_3), 64.08 (d, J_{CH_2-P} = 6.2, O-CH_2-CH_3), 64.12 (3), 56.70 (5"'-O-CH_3), 56.43 (4"'-O-CH_3), 19.20 (d, J_{1-P} = 5.5, 1-Me), 16.04 (d, J_{CH_3-9} = 6.9, O-CH_2-CH_3), 16.00 ppm (d, J_{CH_3-9} = 7.0, O-CH_2-CH_3). ^{31}P NMR (161 MHz, $CDCl_3$, 25 °C): δ = −1.87 ppm. HRMS (ESI+) calculated for $C_{16}H_{25}O_{10}NP$: 422.12106; found: $[M+H]^+$ 422.12143.

4,5-Dimethoxy-2-nitrobenzyl (S)-2-((diethoxyphosphoryl)oxy)-3-methylbutanoate (15). 4,5-Dimethoxy-2-nitrobenzyl (S)-2-hydroxy-3-methylbutanoate **28** (120 mg, 0.38 mmol, 1.00 equiv.) was added to a premixed solution of dry CH_2Cl_2 (3 mL), diethyl chlorophosphate **30** (60 µL, 0.41 mmol, 1.08 equiv.) and NMI (50 µL, 0.63 mmol, 1.66 equiv.) at 25 °C. The mixture was stirred at 25 °C for 12 h. The title compound was isolated (direct injection) by normal-phase flash chromatography (hexanes/ethyl acetate gradient on silica-gel) followed by reversed-phase flash chromatography (water/methanol gradient on C_{18} silica-gel). Yield 15 (29 mg, 17%) of a yellowish oil. ^1H NMR (400 MHz, $CDCl_3$, 25 °C): δ = 7.74 (s, 1H, 3"), 7.13 (s, 1H, 6"), 5.66 (dd, J_{GEM} = 14.9, J_{3a-C} = 0.6, 3a), 5.61 (d, J_{GEM} = 14.9, J_{3b-C} = 0.6, 3b), 4.74 (dd, $J_{1-CH(iPr)}$ = 4.4, J_{1-P} = 7.9, 1), 4.09–4.20 (m, 4H, O-CH_2-CH_3), 4.03 (s, 3H, 5"'-O-CH_3), 3.97 (s, 3H, 4"'-O-CH_3), 2.28 (m, CH^{iPr}), 1.34 (td, 3H, $J_{CH_3-CH_2}$ = 7.1, J_{CH_3-9} = 1.1, O-CH_2-CH_3), 1.31 (td, 3H, $J_{CH_3-CH_2}$ = 7.1, J_{CH_3-P} = 1.1, O-CH_2-CH_3), 1.08 (d, $J_{CH(iPr)-CH(iPr)}$ = 6.9, CH_3^{iPr}), 1.01 ppm (d, $J_{CH_3(iPr)-CH(iPr)}$ = 6.8, CH_3^{iPr}). ^{13}C NMR (100 MHz, $CDCl_3$, 25 °C): δ = 169.45 (d, J_{2-P} = 1.5, 2), 153.81 (5"), 148.26 (4"),

139.63 (2″), 126.63 (1″), 110.32 (6″), 108.09 (3″), 79.89 (d, J_{1-P}= 6.2, 1), 64.22 (d, J_{CH_2-P}=6.1, O–CH_2–CH_3), 64.02 (d, J_{CH_2-P}= 6.2, O–CH_2–CH_3), 63.84 (3), 56.73 (5″–O–CH_3), 56.37 (4″–O–CH_3), 31.69 (d, $J_{CH(Pr)-P}$= 7.0, CH^{iPr}), 18.52 and 16.71 (CH_2^{iPr}), 16.05 (d, J_{CH_3-P}= 7.2, O–CH_2–CH_3), 15.97 ppm (d, J_{CH_3-P}= 7.0, O–CH_2–CH_3). ^{31}P NMR (161 MHz, CDCl$_3$, 25 °C): δ=−1.36 ppm. HRMS (ESI+) calculated for $C_{18}H_{28}O_{10}NNaP$: 472.13430; found: $[M+Na]^+$ 472.13438.

4,5-Dimethoxy-2-nitrobenzyl 2-((diethoxyphosphoryl)oxy)-2-methylpropanoate (16). 4,5-Dimethoxy-2-nitrobenzyl (2-methyl) propanoate 29 (120 mg, 0.40 mmol, 1.00 equiv.) was dissolved in dry CH$_2$Cl$_2$ (3 mL), subsequently adding diethyl chlorophosphate 30 (60 μL, 0.41 mmol, 1.04 equiv.) and NMI (50 μL, 0.63 mmol, 1.57 equiv.) at 25 °C. The mixture was stirred at 25 °C for 36 h and directly injected into a normal-phase flash chromatography system (hexanes/ethyl acetate gradient on silica-gel), followed by reversed-phase flash chromatography (water/methanol gradient on C$_{18}$ silica-gel). Yield 16 (23 mg, 13%) of a yellowish oil. ^{1}H NMR (400 MHz, CDCl$_3$, 25 °C): δ=7.74 (s, 1H, 3″), 7.19 (s, 1H, 6″), 5.63 (s, 2H, 3), 4.12 (m, 4H, O–CH_2–CH_3), 4.03 (s, 3H, 5″–O–CH_3), 3.97 (s, 3H, 4″–O–CH_3), 1.75 (s, 6H, 1-(CH_3)$_2$), 1.32 ppm (td, 6H, J_{CH_3-T}= 7.1, J_{CH_3-P}=1.0, O–CH_2–CH_3). ^{13}C NMR (100 MHz, CDCl$_3$, 25 °C): δ= 171.96 (d, J_{2-P}= 3.9, 2), 153.78 (5″), 148.23 (4″), 139.67 (2″), 126.81 (1″), 110.48 (6″), 108.09 (3″), 80.92 (d, J_{1-P}= 6.2, 1), 64.41 (3), 63.76 (d, J_{CH_2-P}= 6.1, O–CH_2–CH_3), 56.76 (5″–O–CH_3), 56.38 (4″–O–CH_3), 26.30 (d, J_{CH_3-P}= 5.1, 1-(CH_3)$_2$), 16.00 ppm (d, J_{CH_3-P}=7.1, O–CH_2–CH_3). ^{31}P NMR (161 MHz, CDCl$_3$, 25 °C): δ=−4.95 ppm. HRMS (ESI+) calculated for $C_{17}H_{27}O_{10}NP$: 436.13671; found: $[M+H]^+$ 436.13666.

Characterisation of Intermediates (I) and products identified after irradiation with UV light

1-I. ^{13}C NMR (125 MHz, 50% CACO/[D$_6$]DMSO, 25 °C): δ= 174.92 (d, J_{3-P}= 5.1, 3), 160.88 (d, $J_{4'-S}$= 242.6, 4‴), 151.02–151.23 (m, 1′), 147.11–147.31 (m, 1‴), 131.61 (3′), 127.40 (4′), 123.11–123.33 (m, 2‴), 121.37 (d, J_{2-P}= 4.6, 2′), 121.35 (d, J_{2-P}= 4.7, 2′), 118.03 (d, $J_{3'-S}$= 23.7, 3‴), 118.00 (d, $J_{3'-S}$= 23.6, 3‴), 78.25 (d, J_{1-P}= 7.1, 1), 20.79 ppm (d, J_{2-P}= 5.0, 2). ^{31}P NMR (202 MHz, 50% CACO/[D$_6$]DMSO, 25 °C): δ= −12.55 and −12.64 ppm. HRMS (ESI−) calculated for $C_{15}H_{13}O_6FP$: 339.04393; found: $[M−H]^-$ 339.04370.

1-P1, 2-P1. ^{13}C NMR (125 MHz, 50% CACO/[D$_6$]DMSO, 25 °C) 130.76 (3′), 124.66 (4′), 121.56 (d, J_{2-P}= 4.7, 2′), 74.19 (d, J_{1-P}= 6.4, 1), 21.32 ppm (d, J_{2-P}= 2.7, 2). ^{31}P NMR (202 MHz, 50% CACO/[D$_6$]DMSO, 25 °C): δ= −6.17 ppm. HRMS (ESI−) calculated for $C_9H_{10}O_5P$: 245.02205; found: $[M−H]^-$ 245.02221.

1-P, 2-P. ^{13}C NMR (125 MHz, 50% CACO/[D$_6$]DMSO, 25 °C): δ= 180.53 (d, J_{3-P}= 5.5, 3), 73.24 (d, J_{1-P}= 5.5, 1), 21.39 ppm (d, J_{2-P}= 4.3, 2). ^{31}P NMR (202 MHz, 50% CACO/[D$_6$]DMSO, 25 °C): δ= −0.22 ppm. HRMS (ESI−) calculated for $C_3H_8O_6P$: 168.99075; found: $[M−H]^-$ 168.99092.

2-I. ^{13}C NMR (125 MHz, 50% CACO/[D$_6$]DMSO, 25 °C): δ= 175.93 (d, J_{3-P}= 5.4, 3), 151.10–151.27 (m, 1′), 148.83–149.08 (m, 1‴), 136.88–136.99 (m, 4‴), 131.80–131.90 (m, 3″), 131.52–131.60 (m, 3′), 127.31 and 127.29 (4′), 121.29–121.42 (m, 2′), 121.04–121.18 (m, 2″), 78.06 (d, J_{1-P}= 6.6, 1), 21.39–21.46 (m, 4‴–CH_3), 20.83 ppm (d, J_{2-P}= 5.3, 2). ^{31}P NMR (202 MHz, 50% CACO/[D$_6$]DMSO, 25 °C): δ= −12.57 ppm. HRMS (ESI−) calculated for $C_{16}H_{16}O_6P$: 335.06900; found: $[M−H]^-$ 335.06906.

2-P2 ^{13}C NMR (125 MHz, 50% CACO/[D$_6$]DMSO, 25 °C): δ= 179.17 (d, J_{3-P}= 7.3, 3), 151.50 (d, $J_{1'-P}$= 7.0, 1′″), 133.81 (4‴), 131.10 (3‴), 121.30–121.44 (m, 2″), 74.13 (d, J_{1-P}=6.0, 1), 21.41 (d, J_{CH_3-P}= 1.8, 4‴–CH_3), 21.33 ppm (d, J_{2-P}= 3.1, 2). ^{31}P NMR (202 MHz, 50% CACO/[D$_6$]DMSO, 25 °C): δ= −6.01 ppm. HRMS (ESI−) calculated for $C_{10}H_{12}O_6P$: 259.03770; found: $[M−H]^-$ 259.03787.

4-I. ^{13}C NMR (125 MHz, 50% CACO/[D$_6$]DMSO, 25 °C): δ= 123.11–123.33 (m, 2′), 117.90 (d, J_{3-S}= 23.7, 3′), 77.18 (d, J_{1-P}= 6.4, 1), 66.90–67.06 (m, –O–CH_2–CH_3), 20.73–20.92 (m, 2), 16.84–16.99 ppm (m, –O–CH_2–CH_3). ^{31}P NMR (202 MHz, 50% CACO/[D$_6$]DMSO, 25 °C): δ= −7.05 and −7.38 ppm. HRMS (ESI−) calculated for $C_{11}H_{13}O_5FP$: 291.04393; found: $[M−H]^-$ 291.04359.

5-I. ^{13}C NMR (125 MHz, 50% CACO/[D$_6$]DMSO, 25 °C): δ= 176.47 (d, J_{3-P}= 5.6, 3), 175.40 (d, J_{3-P}= 5.8, 3), 151.29 (d, $J_{1'-P}$= 7.0, 1′), 151.25 (d, $J_{1'-P}$= 6.7, 1′), 131.48 and 131.48 (3′), 127.04 and 127.01 (4′), 121.42 (d, J_{2-P}= 4.5, 2′), 121.37 (d, J_{2-P}= 4.6, 2′), 77.08 (d, J_{1-P}= 6.6, 1), 77.06 (d, J_{1-P}= 6.3, 1), 66.85 (d, J_{CH_2-P}= 6.5, –O–CH_2–CH_3), 66.81 (d, J_{CH_2-P}= 6.3, –O–CH_2–CH_3), 20.86 (d, J_{2-P}= 4.8, 2), 20.79 (d, J_{2-P}= 5.1, 2), 16.94 (d, J_{CH_3-P}= 6.7, –O–CH_2–CH_3), 16.91 ppm (d, J_{CH_3-P}= 6.7, –O–CH_2–CH_3). ^{31}P NMR (202 MHz, 50% CACO/[D$_6$]DMSO, 25 °C): δ= −7.14 and −7.53 ppm. HRMS (ESI−) calculated for $C_{13}H_{18}O_6P$: 273.05335; found: $[M−H]^-$ 273.05358.

6-I. ^{13}C NMR (125 MHz, 50% CACO/[D$_6$]DMSO, 25 °C): δ= 176.52 (d, J_{3-P}= 5.4, 3), 176.45 (d, J_{3-P}= 5.8, 3), 149.09 (d, $J_{1'-P}$= 7.1, 1′), 149.06 (d, $J_{1'-P}$= 7.3, 1′), 136.60 and 136.56 (4′), 131.78 and 131.77 (3′), 121.18 (d, J_{2-P}= 4.5, 2′), 121.13 (d, J_{2-P}= 4.6, 2′), 77.02 (d, J_{1-P}= 6.5, 1), 77.00 (d, J_{1-P}= 6.3, 1), 66.77 (d, J_{CH_2-P}= 6.4, –O–CH_2–CH_3), 66.74 (d, J_{CH_2-P}= 6.3, –O–CH_2–CH_3), 21.43 and 21.42 (4′–CH_3), 20.87 (d, J_{2-P}= 4.9, 2), 20.80 (d, J_{2-P}= 5.3, 2), 16.94 (d, J_{CH_3-P}= 6.7, –O–CH_2–CH_3), 16.91 ppm (d, J_{CH_3-P}= 6.4, –O–CH_2–CH_3). ^{31}P NMR (202 MHz, 50% CACO/[D$_6$]DMSO, 25 °C): δ= −6.96 and −7.37 ppm. HRMS (ESI−) calculated for $C_{12}H_{16}O_6P$: 287.06900; found: $[M−H]^-$ 287.06893.

7-I. ^{13}C NMR (125 MHz, 50% CACO/[D$_6$]DMSO, 25 °C): δ= 176.37–176.47 (m, 3), 157.66 and 157.61 (4′), 144.89–145.04 (m, 1′), 122.57 (d, J_{2-P}= 4.4, 2′), 122.52 (d, J_{2-P}= 4.5, 2′), 116.33 and 116.24 (3′), 77.01 (d, J_{1-P}= 6.8, 1), 77.00 (d, J_{1-P}= 6.4, 1), 66.75 (d, J_{CH_2-P}= 6.4, –O–CH_2–CH_3), 66.73 (d, J_{CH_2-P}= 6.4, –O–CH_2–CH_3), 56.91 and 56.84 (4′–O–CH_3), 20.87 (d, J_{2-P}= 5.1, 2), 20.80 (d, J_{2-P}= 5.3, 2), 16.94 (d, J_{CH_3-P}= 6.4, –O–CH_2–CH_3), 16.92 ppm (d, J_{CH_3-P}= 6.3, –O–CH_2–CH_3). ^{31}P NMR (202 MHz, 50% CACO/[D$_6$]DMSO, 25 °C): δ= −6.68, −7.06 ppm. HRMS (ESI−) calculated for $C_{13}H_{18}O_7P$: 303.06391; found: $[M−H]^-$ 303.06411.

3-P2, 4-P2, 5-P2, 6-P2 and 7-P2. ^{13}C NMR (125 MHz, 50% CACO/[D$_6$]DMSO, 25 °C): δ= 179.87 (d, J_{3-P}= 6.5, 3), 73.52 (d, J_{1-P}= 5.4, 1), 62.42 (d, J_{CH_2-P}= 5.9, –O–CH_2–CH_3), 21.42 (d, J_{2-P}= 3.2, 2), 17.33 ppm (d, J_{CH_2-P}= 7.5, –O–CH_2–CH_3). ^{31}P NMR (202 MHz, 50% CACO/[D$_6$]DMSO, 25 °C): δ= −1.04 ppm. HRMS (ESI−) calculated for $C_6H_{12}O_6P$: 197.02205; found: $[M−H]^-$ 197.02187.

8-I. ^{13}C NMR (125 MHz, 50% CACO/[D$_6$]DMSO, 25 °C): δ= 175.36 (d, J_{3-P}= 2.8, 2), 175.29 (d, J_{3-P}= 2.8, 4), 151.31–151.53 (m, 1′), 131.40 (3′), 126.93 and 126.87 (4′), 121.45 (d, J_{2-P}= 4.4, 2′), 121.38 (d, J_{2-P}= 4.4, 2′), 85.66 (d, J_{1-P}= 7.0, 1), 85.59 (d, J_{1-P}= 6.2, 1), 66.66–66.84 (m, –O–CH_2–CH_3), 32.29–32.46 (m, CH^{iPr}), 20.11 and 20.08 and 18.04 and 17.86 (CH_2^{iPr}), 16.95 ppm (d, J_{CH_3-P}= 6.8, –O–CH_2–CH_3). ^{31}P NMR (202 MHz, 50% CACO/[D$_6$]DMSO, 25 °C): δ= −6.64 and −7.09 ppm. HRMS (ESI−) calculated for $C_{13}H_{18}O_6P$: 301.08465; found: $[M−H]^-$ 301.08489.

8-P2. ^{13}C NMR (125 MHz, 50% CACO/[D$_6$]DMSO, 25 °C): δ= 178.68 (d, J_{3-P}= 3.1, 2), 82.29 (d, J_{1-P}= 5.9, 1), 62.38 (d, J_{CH_2-P}= 5.8, –O–CH_2–CH_3), 32.73 (d, $J_{CH(iPr)-P}$= 5.2, CH^{iPr}), 19.97 and 18.93 (CH_2^{iPr}), 17.37 ppm (d, J_{CH_3-P}= 8.0, –O–CH_2–CH_3). ^{31}P NMR (202 MHz, 50% CACO/[D$_6$]DMSO, 25 °C): δ= −0.60 ppm. HRMS (ESI−) calculated for $C_7H_{14}O_6P$: 225.05335; found: $[M−H]^-$ 225.05350.

9-P2. ^{13}C NMR (125 MHz, 50% CACO/[D$_6$]DMSO, 25 °C): δ= 181.68 (2), 81.23 (d, J_{1-P}= 7.9, 1), 62.21 (d, J_{CH_2-P}= 5.5, –O–CH_2–CH_3), 27.32 (d, J_{CH_3-P}= 1.8, 1-(CH_3)$_2$), 17.33 ppm (d, J_{CH_3-P}= 7.7, –O–CH_2–CH_3). ^{31}P NMR (202 MHz, 50% CACO/[D$_6$]DMSO, 25 °C): δ= −4.22 ppm. HRMS

Chemistry–A European Journal

Full Paper
doi.org/10.1002/chem.202101805

Chemistry
Europe
European Chemical
Societies Publishing

(ESI−) calculated for $C_6H_{12}O_6P$: 211.03770; found: [*M*−H]⁻ 211.03777.

10-P1. ¹³C NMR (125 MHz, 25% CACO/[D₆]DMSO, 25 °C): δ= 129.82 (3'), 123.01 (4'), 120.84 (d, $J_{2'-P}$= 4.9, 2'), 82.20 (d, J_{1-P}= 6.5, 1), 32.05 (d, $J_{CH_2P_1-P}$= 4.6, CH^P), 19.92 and 18.47 ppm (CH₃^P). C2 and C1' were not detected. ³¹P NMR (202 MHz, 25% CACO/[D₆]DMSO, 25 °C): δ= −5.91 ppm. HRMS (ESI−) calculated for $C_{11}H_{16}O_6P$: 273.05335; found: [*M*−H]⁻ 273.05330.

10-P and 11-P. ¹³C NMR (125 MHz, 25% CACO/[D₆]DMSO, 25 °C): δ= 178.38 (2), 81.02 (d, J_{1-P}= 6.5, 1), 31.65 (d, $J_{CH_2P_1-P}$= 7.6, CH^P), 20.39 and 17.74 ppm (CH₃^P). ³¹P NMR (202 MHz, 50% CACO/[D₆]DMSO, 25 °C): δ= −0.99 ppm. HRMS (ESI−) calculated for $C_5H_{10}O_6P$: 197.02205; found: [*M*−H]⁻ 197.02181.

11-I. ¹³C NMR (125 MHz, 25% CACO/[D₆]DMSO, 25 °C): δ= 172.80 (m, 2), 148.82–149.01 (m, 1'''), 135.57 (4'''), 131.12 and 131.03 and 130.84 and 130.75 (3''', 3'), 126.25 (4'), 121.13 (d, $J_{2'2''-P}$= 4.8, 2' or 2'''), 120.90 (d, $J_{2'2''-P}$= 4.5, 2' or 2'''), 120.83 (d, $J_{2'2''-P}$= 4.6, 2' or 2'''), 120.55–120.67 (m, 2' or 2'''), 86.13–86.25 (m, 1), 31.74 (d, $J_{CH_2P_1-P}$= 7.2, CH^P), 21.06 and 21.04 (4'''-CH₃), 19.99 and 19.97 and 17.54 and 17.53 ppm (CH₃^P). ³¹P NMR (202 MHz, 25% CACO/[D₆]DMSO, 25 °C): δ= −12.16 and −12.19 ppm. HRMS (ESI−) calculated for $C_{18}H_{20}O_6P$: 363.10030; found: [*M*−H]⁻ 363.10056.

11-P2. ¹³C NMR (125 MHz, 25% CACO/[D₆]DMSO, 25 °C): δ= 176.85 (d, J_{2-P}= 3.1, 2), 151.92 (d, J_{1-P}= 6.4, 1''), 131.79 (4''), 130.16 (3''), 120.63 (d, $J_{2'-P}$= 4.9, 2''), 82.11 (d, J_{1-P}= 6.9, 1), 32.05 (d, $J_{CH_2P_1-P}$= 5.2, CH^P), 21.04 (4'''-CH₃), 19.94 and 18.48 ppm (CH₃^P). ³¹P NMR (202 MHz, 25% CACO/[D₆]DMSO, 25 °C): δ= −5.76 ppm. HRMS (ESI−) calculated for $C_{12}H_{16}O_6P$: 287.06900; found: [*M*−H]⁻ 287.06927.

12-P and 13-P. ¹³C NMR (125 MHz, 50% CACO/[D₆]DMSO, 25 °C): δ= 182.51–182.64 (m, 2), 81.41 (d, J_{1-P}= 7.5, 1), 28.05 ppm (d, J_{CH_2-P}= 3.7, 1-(CH₂)₂). ³¹P NMR (202 MHz, 50% CACO/[D₆]DMSO, 25 °C): δ= −2.68 ppm. HRMS (ESI−) calculated for $C_4H_8O_6P$: 183.00640; found: [*M*−H]⁻ 183.00654.

13-P2. ¹³C NMR (125 MHz, 50% CACO/[D₆]DMSO, 25 °C): δ= 133.48 (4'''), 131.04 (3'''), 121.16 (d, $J_{2''-P}$= 4.7, 2''), 81.97 (d, J_{1-P}= 7.9, 1), 27.32 (d, J_{CH_2-P}= 2.2, 1-(CH₂)₂), 21.40 ppm (4'''-CH₃). ³¹P NMR (202 MHz, 50% CACO/[D₆]DMSO, 25 °C): δ= −9.30 ppm. HRMS (ESI +) calculated for $C_{11}H_{16}O_6P$: 273.05335; found: [*M*−H]⁻ 273.05389.

14-I. ¹³C NMR (125 MHz, 50% CACO/[D₆]DMSO, 25 °C): δ= 176.89 (J_{1-P}= 5.5, 3), 76.89 (d, J_{1-P}= 6.0, 1), 65.76 (d, J_{T-P}= 5.6, 1'), 65.71 (d, J_{T-P}= 5.4, 1'), 20.90 (d, J_{2-P}= 5.0, 2), 16.97 (d, J_{T-P}= 6.6, 2'), 16.96 ppm (d, J_{T-P}= 6.4, 2'). ³¹P NMR (202 MHz, 50% CACO/[D₆]DMSO, 25 °C): δ= −2.01 ppm. HRMS (ESI−) calculated for $C_7H_{14}O_6P$: 225.05335; found: [*M*−H]⁻ 225.05355.

15-I. ¹³C NMR (125 MHz, 50% CACO/[D₆]DMSO, 25 °C): δ= 175.78 (d, J_{2-P}= 1.8, 2), 84.72 (d, J_{1-P}= 6.8, 1), 65.65–65.75 (m, 1'), 32.36 (d, $J_{CH_2P_1-P}$= 6.8, CH^P), 20.14 (CH₂^P), 18.08 (CH₂^P), 17.01 (d, J_{T-P}= 6.3, 2'), 16.99 ppm (d, J_{T-P}= 6.4, 2'). ³¹P NMR (202 MHz, 50% CACO/[D₆]DMSO, 25 °C): δ= −1.52 ppm. HRMS (ESI−) calculated for $C_9H_{18}O_6P$: 253.08465; found: [*M*−H]⁻ 253.08479.

16-I. ¹³C NMR (125 MHz, 50% CACO/[D₆]DMSO, 25 °C): δ= 178.58 (d, J_{2-P}= 6.6, 2), 86.18 (d, J_{1-P}= 7.4, 1), 65.31 (d, J_{T-P}= 6.1, 1'), 27.53 (d, J_{CH_2-P}= 3.6, 1-(CH₂)₂), 16.96 ppm (d, J_{T-P}= 6.9, 2'). ³¹P NMR (202 MHz, 50% CACO/[D₆]DMSO, 25 °C): δ= −5.44 ppm. HRMS (ESI−) calculated for $C_6H_{14}O_6P$: 239.06900; found: [*M*−H]⁻ 239.06903.

NMR spectroscopy. For compound characterisation, NMR spectra were recorded on a Bruker Avance III spectrometer operating at 400 MHz for ¹H, 100 MHz for ¹³C and 161 MHz for ³¹P. Irradiation NMR experiments were performed on a Bruker Avance III spectrometer with a broad-band cryo probe with an ATM module (5 mm CPBBO BB-¹H/¹⁹F/¹⁵N/D Z-GRD) operating at 500 MHz for ¹H,

125.7 MHz for ¹³C and 202 MHz for ³¹P. For NMR signal assignment, standard Bruker pulse sequences were employed for both 1D (¹H, ¹³C-APT, ³¹P) and 2D (COSY, ROESY, HSQC, HMBC) NMR experiments at a corrected temperature of 25 °C. The structure of each intermediate was determined *in situ*, as recently reported.[17] All NMR data was interpreted using Topspin 3.5. For reference, the following solvent signals were used: [D₆]DMSO: 2.50 (¹H) and 39.5 (¹³C) ppm; CDCl₃: 7.27 (¹H) and 77.0 (¹³C). ³¹P NMR spectra were referenced to H₃PO₄ with 0 ppm.

For NMR experiments with in situ irradiation, a light emitting diode (LED; Thorlabs, Germany) was used at 365 nm. The light was guided into the spectrometer, directly into the NMR tube, via a multimode silica optical fiber with 1-mm diameter, 0.39 NA, and a high amount of OH (Thorlabs, Germany). The irradiation setup was described in detail in our previous work.[18] For in situ irradiation NMR experiments, we used 0.5 mM solutions of the corresponding linkers 1–13 in a mixture of cacodylate buffer (CACO) with [D₆]DMSO (1:1, v/v or, in some case, 1:3, v/v). All irradiation experiments were performed at room temperature.

Mass spectrometry. Mass spectra were measured on an LTQ Orbitrap XL (Thermo Fisher Scientific) by electrospray ionisation (ESI).

Conflict of Interest

The authors declare no conflict of interest.

[1] A. Alouane, R. Labruère, T. Le Saux, F. Schmidt, L. Jullien, *Angew. Chem. Int. Ed.* 2015, *54*, 7492–7509; *Angew. Chem.* 2015, *127*, 7600–7619.

[2] a) R. V. Gonzaga, L. A. do Nascimento, S. S. Santos, B. A. Macha-do Sanches, J. Giarolla, E. I. Ferreira, *J. Pharm. Sci.* 2020, *109*, 3262–3281; b) Y. Huang, Y. Jiang, H. Wang, J. Wang, M. C. Shin, Y. Byun, H. He, Y. Liang, V. C. Yang, *Adv. Drug Delivery Rev.* 2013, *65*, 1299–1315; c) A. Dal Corso, L. Pignataro, L. Belvisi, C. Gennari, *Chem. Eur. J.* 2019, *25*, 14740–14757.

[3] a) X. Yang, Z. Pan, M. R. Choudhury, Z. Yuan, A. Anifowose, B. Yu, W. Wang, B. Wang, *Med. Res. Rev.* 2020, *40*, 2682–2713; b) A. D. Corso, V. Borlandelli, C. Corno, P. Perego, L. Belvisi, L. Pignataro, C. Gennari, *Angew. Chem. Int. Ed.* 2020, *59*, 4176–4181; *Angew. Chem.* 2020, *132*, 4206–4210.

[4] J. C. Kern, D. Dooney, R. Zhang, L. Liang, P. E. Brandish, M. Cheng, G. Feng, A. Beck, D. Bresson, J. Firdos, D. Gately, N. Knudsen, A. Manibusan, Y. Sun, R. M. Garbaccio, *Bioconjugate Chem.* 2016, *27*, 2081–2088.

[5] a) J. Yan, S. Lee, A. Zhang, J. Yoon, *Chem. Soc. Rev.* 2018, *47*, 6900–6916; b) R. Labruère, A. Alouane, T. Le Saux, I. Aujard, P. Pelupessy, A. Gautier, S. Dubruille, F. Schmidt, L. Jullien, *Angew. Chem. Int. Ed.* 2012, *51*, 9344–9347; *Angew. Chem.* 2012, *124*, 9478–9481.

[6] D. A. Roberts, B. S. Pilgrim, T. N. Dell, M. M. Stevens, *Chem. Sci.* 2020, *11*, 3713–3718.

Chemistry–A European Journal | Full Paper
doi.org/10.1002/chem.202101805

**Chemistry
Europe**
European Chemical
Societies Publishing

[7] a) R. J. Amir, N. Pessah, M. Shamis, D. Shabat, *Angew. Chem. Int. Ed.* 2003, *42*, 4632–4637; *Angew. Chem.* 2003, *115*, 4632–4637; b) G. Liu, X. Wang, J. Hu, G. Zhang, S. Liu, *J. Am. Chem. Soc.* 2014, *136*, 7492–7497.

[8] A. Sagi, R. Weinstain, N. Karton, D. Shabat, *J. Am. Chem. Soc.* 2008, *130*, 5434–5435.

[9] H. Wang, R. Wang, K. Cai, H. He, Y. Liu, J. Yen, Z. Wang, M. Xu, Y. Sun, X. Zhou, Q. Yin, L. Tang, I. T. Dobrucki, L. W. Dobrucki, E. J. Chaney, S. A. Boppart, T. M. Fan, S. Lezmi, X. Chen, L. Yin, J. Cheng, *Nat. Chem. Biol.* 2017, *13*, 415–424.

[10] A. J. Burt, P. Ahmadvand, L. K. Opp, A. T. Ryan, C. Kang, R. J. Mancini, *ChemMedChem* 2020, *15*, 2004–2009.

[11] a) T. Gollnest, T. D. De Oliveira, D. Schols, J. Balzarini, C. Meier, *Nat. Commun.* 2015, *6*, 1–15; b) Y. Wei, G. Qiu, B. Lei, L. Qin, H. Chu, Y. Lu, G. Zhu, Q. Gao, Q. Huang, G. Qian, P. Liao, X. Luo, X. Zhang, C. Zhang, Y. Li, S. Zhang, Y. Yu, P. Tang, J. Ni, P. Yan, Y. Zhou, P. Li, X. Huang, A. Gong, J. Liu, *J. Med. Chem.* 2017, *60*, 8580–8590; c) C. J. Choy, C. R. Ley, A. L. Davis, B. S. Backer, J. J. Geruntho, B. H. Clowers, C. E. Berkman, *Bioconjugate Chem.* 2016, *27*, 2206–2213; d) F. P. Olatunji, J. W. Herman, B. N. Kesic, D. Olabode, C. E. Berkman, *Tetrahedron Lett.* 2020, *61*, 152398.

[12] S. Gnaim, D. Shabat, *Acc. Chem. Res.* 2019, *52*, 2806–2817.

[13] M. R. Levengood, X. Zhang, J. H. Hunter, K. K. Emmerton, J. B. Miyamoto, T. S. Lewis, P. D. Senter, *Angew. Chem. Int. Ed.* 2017, *56*, 733–737; *Angew. Chem.* 2017, *129*, 751–755.

[14] a) M. A. Fanale, S. M. Horwitz, A. Forero-Torres, N. L. Bartlett, R. H. Advani, B. Pro, R. W. Chen, A. Davies, T. Illidge, D. Huebner, D. A. Kennedy, A. R. Shustov, *J. Clin. Oncol.* 2014, *32*, 3137–3143; b) A. Younes, J. M. Connors, S. I. Park, M. Fanale, M. M. O'Meara, N. N. Hunder, D. Huebner, S. M. Ansell, *Lancet Oncol.* 2013, *14*, 1348–1356.

[15] a) A. N. Tevyashova, A. M. Korolev, A. S. Trenin, L. G. Dezhenkova, A. A. Shtil, V. I. Polshakov, O. Y. Savelyev, E. N. Olsufyeva, *J. Antibiot. (Tokyo).* 2016, *69*, 549–560; b) J. Cui, B. Ren, Y. Tong, H. Dai, L. Zhang, *Virulence* 2015, *6*, 362–371; c) L. Tsenova, K. Sokol, V. H. Freedman, G. Kaplan, *J. Infect. Dis.* 2008, *177*, 1563–1572.

[16] A. F. L. Schneider, C. P. R. Hackenberger, *Curr. Opin. Biotechnol.* 2017, *48*, 61–68.

[17] E. Procházková, R. Navrátil, Z. Janeba, J. Roithová, O. Baszczyński, *Org. Biomol. Chem.* 2019, *17*, 315–320.

[18] E. Procházková, P. Šimon, M. Straka, J. Filo, M. Majek, M. Cigáň, O. Baszczyński, *Chem. Commun.* 2020, *1*, 14–17.

[19] a) R. M. Beesley, C. K. Ingold, J. F. Thorpe, *J. Chem. Soc. Trans.* 1915, *107*, 1080–1106; b) M. E. Jung, G. Piizzi, *Chem. Rev.* 2005, *105*, 1735–1766.

[20] E. Procházková, J. Filo, M. Cigáň, O. Baszczyński, *Eur. J. Org. Chem.* 2020, *7*, 897–906.

[21] A. P. Esser-Kahn, N. R. Sottos, S. R. White, J. S. Moore, *J. Am. Chem. Soc.* 2010, *132*, 10266–10268.

[22] P. Gao, X. Nie, M. Zou, Y. Shi, G. Cheng, *J. Antibiot. (Tokyo).* 2011, *64*, 625–634.

[23] a) M. Fiore, *Org. Biomol. Chem.* 2018, *16*, 3068–3086; b) A. Miccoli, B. A. Dhiani, P. J. Thornton, O. A. Lambourne, E. James, H. Kadri, Y. Mehellou, *ChemMedChem* 2020, *15*, 671–674.

[24] a) P. Klán, T. Šolomek, C. G. Bochet, A. Blanc, R. Givens, M. Rubina, V. Popik, A. Kostikov, J. Wirz, *Chem. Rev.* 2013, *113*, 119–191; b) I. Aujard, C. Benbrahim, M. Gouget, O. Ruel, J. B. Baudin, P. Neveu, L. Jullien, *Chem. Eur. J.* 2006, *12*, 6865–6879.

[25] a) V. M. Dhurandhare, G. P. Mishra, S. Lam, C. C. Wang, *Org. Biomol. Chem.* 2015, *13*, 9457–9461; b) E. Granger, K. Solomianko, C. Young, J. Erb, *Tetrahedron Lett.* 2018, *59*, 1404–1408.

[26] U. Pradere, E. C. Garnier-Amblard, S. J. Coats, F. Amblard, R. F. Schinazi, *Chem. Rev.* 2014, *114*, 9154–9218.

[27] a) K. Roy, P. L. A. Popelier, *J. Phys. Org. Chem.* 2009, *22*, 186–196; b) M. D. Liptak, K. C. Gross, P. G. Seybold, S. Feldgus, G. C. Shields, *J. Am. Chem. Soc.* 2002, *124*, 6421–6427.

[28] I. Sharma, G. A. Kaminski, *J. Comput. Chem.* 2012, *33*, 2388–2399.

[29] G. Chuchani, A. Frohlich, *J. Chem. Soc. B Phys. Org.* 1971, *78*, 1417–1420.

[30] L. Čechová, J. Kind, M. Dračínský, J. Filo, Z. Janeba, C. M. Thiele, M. Cigáň, E. Procházková, *J. Org. Chem.* 2018, *83*, 5986–5998.

II.

M. Đud, **M. Tichotová**, E. Procházková and O. Baszczyňski:

Phosphate-Based Self-Immolative Linkers for the Delivery of Amine-Containing Drugs

Article

Phosphate-Based Self-Immolative Linkers for the Delivery of Amine-Containing Drugs

Mateja Đud [1,†], Markéta Tichotová [1,2,†], Eliška Procházková [2,*] and Ondřej Baszczyňski [1,2,*]

[1] Faculty of Science, Charles University, 128 43 Prague, Czech Republic; mateja.dud@natur.cuni.cz (M.Đ.); marketa.tichotova@uochb.cas.cz (M.T.)

[2] Institute of Organic Chemistry and Biochemistry, The Czech Academy of Sciences, 166 10 Prague, Czech Republic

[*] Correspondence: prochazkova@uochb.cas.cz (E.P.); ondrej.baszczynski@natur.cuni.cz (O.B.)

[†] These authors contributed equally to this work.

Abstract: Amine-containing drugs often show poor pharmacological properties, but these disadvantages can be overcome by using a prodrug approach involving self-immolative linkers. Accordingly, we designed L-lactate linkers as ideal candidates for amine delivery. Furthermore, we designed linkers bearing two different cargos (aniline and phenol) for preferential amine cargo release within 15 min. Since the linkers carrying secondary amine cargo showed high stability at physiological pH, we used our strategy to prepare phosphate-based prodrugs of the antibiotic Ciprofloxacin. Therefore, our study will facilitate the rational design of new and more effective drug delivery systems for amine-containing drugs.

1. Introduction

Drugs containing an amino group are key pharmaceutical agents, covering a broad spectrum of biological actions and displaying anti-inflammatory [1], anticancer [2,3], antimicrobial [4,5], and pain-relieving properties [6]. Currently, 542 drugs containing an amino group have already been approved for the EU market, according to the drug bank online (https://go.drugbank.com, accessed on 15 July 2021). This number does not include many other biologically active compounds—potential leads—or compounds from natural resources, such as alkaloids. However, amine drugs often show poor pharmacological properties, such as low aqueous solubility and poor membrane permeability due to ionization of the amino group [7] under physiological conditions. Nevertheless, amines are generally susceptible to derivatization. Thus, a prodrug strategy [8–10] can be used to overcome these drawbacks [7,11].

Prodrug strategies rely on a structural modification (masking) of the active pharmaceutical agent—a drug—with a suitable protecting group (promoiety) to modulate its pharmacokinetic properties. Such a change helps to facilitate drug delivery to the target site (e.g., tissues, cells, cell compartments, or organs). One of the most rapidly developing prodrug strategies consists of using self-immolative (SI) linkers [12,13] to control drug release [14].

SI linkers are covalent assemblies that couple an active compound (drug) to a protecting group. After external stimuli, either chemical or enzymatic, a cascade of spontaneous reactions [15] leads to linker fragmentation and consequently to drug release. The two main classes of SI linkers are (1) carbamates and (2) phosphate-based systems. Phosphorus-based SI linkers stand above the "classical" carbamate linkers because they make it possible to attach an additional substituent, which can help fine-tune the SI rate or provide a double cargo option.

Phosphorus-based SI linkers have been introduced as suitable drug-delivery systems for several drugs. A paradigmatic example of a phosphorus-based SI linker application is the methoxymethylphosphonic acid (MMPA) drug delivery vehicle for the oral delivery of propofol [16]. Other examples include pro-nucleotide prodrugs (ProTides) [17], which have been used to treat various viral infections, such as HIV [18], hepatitis B [19], or SARS-CoV-2 (COVID-19) [20]. Considering their success as drug-delivery systems, phosphate-based SI linkers may find broader applications in drug discovery and materials science through systematic studies.

This study reports the development of new phosphate-based SI linkers designed to release amine-containing cargos (Figure 1). For this purpose, we searched for a suitable spacer responsible for SI. Although SI of ethylene glycol linkers 1–6 did not lead to amine release in a reasonable time, our lactate linkers 7–9 showed successful cargo release. After screening a wide range of linkers bearing various amines (10–16), representing model drugs, we prepared a prodrug of the FDA-approved drug Ciprofloxacin, which is a broad-spectrum fluoroquinolone antibiotic.

Figure 1. Self-immolation (SI) of 1–16; (a) SI is initiated by UV light (365 nm), which cleaves a photosensitive DMNB group [21]; (b) the intermediate I then spontaneously cyclizes, releasing R₁ (4–6) or LG (7–16); (c) hydrolysis of the lactate-based cyclic intermediate leads to the final product P, whereas glycol-based **cyc-I** hydrolyzes to the final product **P2** or the cyclic product **cyc-P**.

2. Results

2.1. Ethylene Glycol Phosphate-Based Linkers

Ethylene glycol-based linkers provide stable cyclic intermediates traceable by NMR spectroscopy, which indicates that the cargo is released via SI [22]. For proof of concept, we prepared glycol-based linkers 1–6 (Scheme 1) bearing phenethylamine, piperidine, and aniline (LG-1, LG-2 and LG-3, respectively) as representatives of primary, secondary, and aromatic amines, respectively.

Target linkers were synthesized in two consecutive phosphorylation steps, starting from commercially available phosphorodichloridates 20–21 (Scheme 1). Carbonate 19 was prepared in a reaction between ethylene glycol and dimethoxynitrobenzyl (DMNB) chloroformate in THF, using pyridine as the base [22]. The reaction of equimolar amounts of dichloridates 20–21 and DMNB carbonate 19 in the presence of triethylamine (TEA) in toluene gave intermediates 22 and 23. This reaction was monitored by ^{31}P-NMR spectra (signal at δ_P 4.7 ppm for 22 and δ_P 0.08 ppm for 23, in CDCl$_3$), and intermediates 22 and 23 were directly used for the second phosphorylation step. The reaction of 22 and 23 with one equivalent of the corresponding amine and TEA as a base afforded the final linkers 1–6. The isolated yields in the series bearing ethoxy substituent (1–3) were moderate (23–44%), whereas compounds 4–6 were isolated in low yields (4–12%) despite multiple flash chromatography (silica gel and C18).
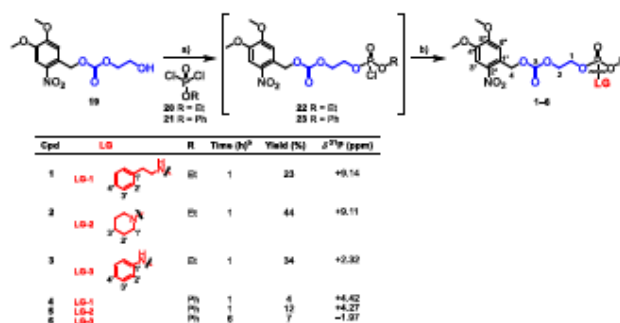
Scheme 1. Linkers 1–6 bearing ethylene glycol spacer were synthesized under the following reaction conditions: (a) ethyl (**20**) or phenyl (**21**) dichlorophosphate (0.5 mmol), TEA (0.65 mmol), toluene, 25 °C, 16 h; (b) the corresponding amine (0.5 mmol), TEA (0.5 mmol), toluene, 25 °C, 1–6 h.

The self-immolation of 1–6 was triggered photochemically (365 nm), and the reaction was monitored by ^{31}P-NMR spectroscopy. Surprisingly, the successful photoactivation of 1–3 yielded intermediates (1–3)-**I**, with no further spectral change over several days, thus indicating that the cargo was not released (Figure S1 in Supplementary Materials).

In contrast to 1–3, linkers 4 and 6 afforded cyclic intermediates **4-cyc-I** and **6-cyc-I** within 5 min of irradiation (Figure 2). In compound 5, bearing a secondary amine, only a trace of **5-cyc-I** was detected overnight. However, the downfield shifted ^{31}P-NMR signals of cyclic intermediates **4-cyc-I**, **5-cyc-I**, and **6-cyc-I** (δ_P 28.0, 26.5, 21.7 ppm, respectively) indicated that the amine cargo was still attached to the phosphorus and that phenol was released instead.

^{31}P-NMR spectra recorded overnight suggested three different scenarios: (1) preferential release of phenol as **4-cyc-I** and **4-P2** were detected; (2) formation of the stable intermediate **5-I**, which released phenol in several days; (3) sequential release of phenol in minutes (**6-cyc-I**) and aniline overnight (**6-cyc-P**). The formation of phospholane intermediates (4-6)-**cyc-I** could be useful for other applications, such as preparing functional biopolymers for controlled drug delivery systems [23], but linkers 1–6 did not release amine successfully. Therefore, we altered the spacer structure to find an effective system for releasing amine-containing cargos.
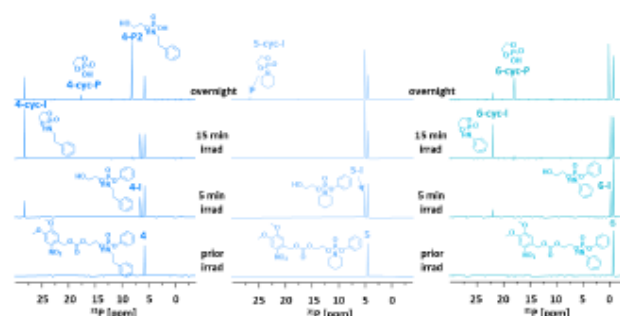

Figure 2. ^{31}P-NMR spectra of linkers 4–6 (5 mM in 50% CACO/DMSO) measured before and after irradiation with UV light (365 nm) at room temperature (25 °C), previously optimizing the reaction conditions (DMSO/cacodylate buffer (CACO), 0.1 M, pH = 7.4; 1/1, *v/v*) [24].

2.2. Lactate Phosphate-Based Linkers

To stimulate amine cargo release, we altered the glycol spacer in **1–3** to an L-lactate spacer, thus preparing linkers **7–9** (Scheme 2). Despite promoting a slow release of phenolic compounds, [25] the lactate spacer could be suitable for amines. Compounds **7–9** were synthesized from DMNB ester **24** via acid-catalyzed esterification in refluxing toluene [25], and intermediate **25** was generated in a reaction of **24** with dichloridate **20** in toluene. The reaction was monitored by ^{31}P-NMR, and a new pair of ^{31}P-NMR signals at δ_P 4.6 ppm and δ_P 4.1 ppm (ca. 1:1 ratio), corresponding to two diastereoisomers of **25**, was observed due to the new stereogenic center on phosphorus. Intermediate **25** was directly used for amine phosphorylation, and the final products **7–9** were isolated with good yields (42–52%) as 1:1 mixtures of diastereoisomers, as shown by the two sets of NMR signals.

Cpd	LG	Yield (%)	δ^{31}P (ppm)
7	LG-1	52	+9.07; +8.40
8	LG-2	42	+8.81; +8.30
9	LG-3	46	+2.15; +1.87
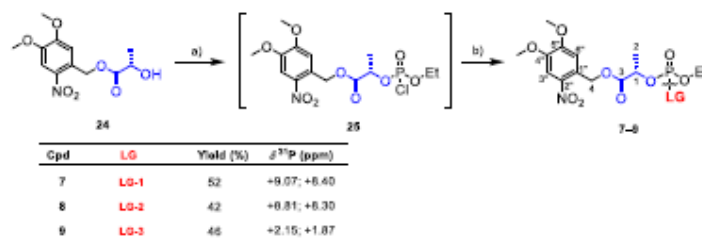
Scheme 2. Lactate-based linkers **7–9** were synthesized under the following reaction conditions: (a) ethyl dichlorophosphate **20** (0.5 mmol), TEA (0.65 mmol), dry toluene, 25 °C, 16 h; (b) the corresponding amine (0.5 mmol), TEA (0.5 mmol), dry toluene, 25 °C, 1 h.

The SI of **7–9**, monitored by ^{31}P-NMR, showed successful amine release, which provided the final product **P** in 15 min (Figure S2 in the Supplementary Materials). However, in **7** and **9**, we also detected a new ^{31}P signal (δ_P −1.6 ppm) belonging to the undesired product of alternative decomposition (**7-X** and **9-X**). Interestingly, linker **8**, bearing a secondary amine as a cargo, did not form the undesired product **8-X** and followed the expected reaction course.

Given the unexpected reactivity of **7** and **9** in the CACO/DMSO mixture (1/1, v/v), we optimized the solvent system. For this purpose, we performed irradiation experiments on **7**, in various solvent mixtures (Figure S3 in the Supplementary Materials), and we found that the formation of side product **7-X** can be suppressed by either decreasing the pH of the cacodylate buffer (to pH = 5) or changing the buffer itself. HEPES buffer or an unbuffered system can suppress the formation of **7-X**. Lastly, we selected the HEPES (pH 7.4)/DMSO system (1:1, v/v) for further SI investigation.

We monitored the SI of **7–9** in the HEPES/DMSO solvent mixture (Figure 3). In 5 min, we detected the final product **P** in all cases (δ_P −1.1 ppm). Linkers **7** and **8** did not provide any intermediate **7-I** and **8-I**, respectively, as they undergo fast cyclization, and photoactivation is a rate-limiting step. In turn, the SI of **9** was slow, and we did detect intermediate **9-I** (Figure 3, right) or traces of the undesired product **9-X** (δ_P −1.6 ppm). Considering the overall limited stability of **7** and **9**, we investigated the formation of the unknown side product **X** and performed stability tests in **7–9**. The results showed the limited stability of **7** and **9**, which were significantly decomposed within 7 days (Figures S6 and S7 in the Supplementary Materials, details therein).

2.3. Characterization of the Undesired Product X

To identify the alternative decomposition pathway, we characterized the undesired product **X**. The significant upfield shift (δ_P < 0) suggested that the P-NH bond had been cleaved. In addition, one singlet ^{31}P-NMR signal indicated the lack of a stereogenic center on the phosphorus atom. ^{31}P-NMR chemical shifts of **7-X** and **9-X** differed slightly

(δ_P −1.66 and −1.58 ppm, respectively), as found in the starting compounds **7** and **9** (differing in LGs). Combined, these findings demonstrate that **7-X** and **9-X** also differ in the amine moiety, which may be explained by the intramolecular rearrangement proposed in Figure 4. Additional 2D NMR experiments performed on linker **7** suggested the formation of carboxamide **7-X**. We found the key interaction between the phenethylamine alkyl chain and lactate carbonyl in the HMBC spectrum (Figure 4). The proposed structure of **7-X** was confirmed by HR-MS (Figure S64 in the Supplementary Materials).
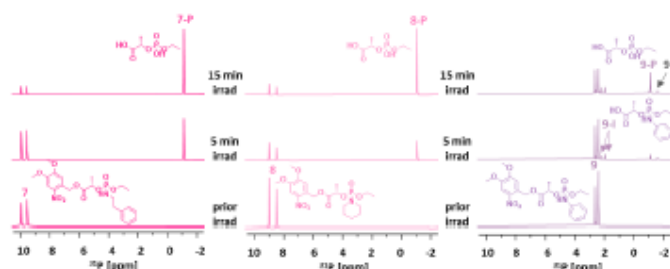
Figure 3. ^{31}P NMR spectra of linkers **7–9** (5 mM), measured before and after irradiation by UV light (365 nm) in a solvent mixture of 50% HEPES (pH = 7.4)/DMSO at room temperature (25 °C).
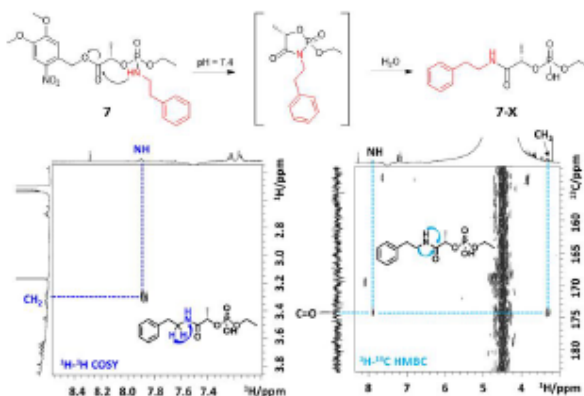
Figure 4. The proposed mechanism of intramolecular rearrangement: part of the COSY and HMBC spectra recorded after UV irradiation of **7** (5 mM solution in CACO (pH 7.4)/DMSO) show cross-peaks between lactate C=O and phenethylamine NH-CH$_2$ groups, confirming the structure of **7-X**.

Only the linkers containing the NH group in the phosphoramidate bond underwent the proposed intramolecular rearrangement (linker **8** without NH did not form **8-X**). Indeed, *N*-methylation of **9** yielded linker **10**, which did not form the undesired product **10-X** (Figure S4 in the Supplementary Materials). This intramolecular rearrangement has already been reported by the Mulliez group in 1985 [26].

2.4. Amine Screening—Application Scope

Based on the successful SI observed in linkers **7–9**, we examined the synthetic scope and application feasibility of lactate linkers by altering the structure of the cargo. Accordingly, we prepared linkers **10–16** (Scheme 3).

Cpd	LG	R	Time (h)[b]	Yield (%)	$\delta^{31}P$ (ppm)
10		Et	24	18	+5.83; +5.55
11		Et	1	46	+7.70; +7.16
12		Et	1	42	+8.61; +7.87
13		Et	1	29	−1.00; −1.20
14	LG-1	Ph	1	29	+4.40; +3.55
15	LG-2	Ph	1	30	+3.96; +3.47
16	LG-3	Ph	2	21	−2.50; −2.99
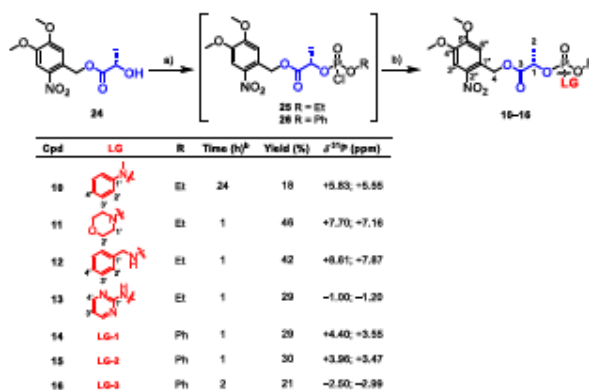
Scheme 3. Lactate-based linkers **10–16** were synthesized under the following reaction conditions: (a) ethyl or phenyl dichlorophosphate **20–21** (0.5 mmol), TEA (0.65 mmol), dry toluene, 25 °C, 16 h; (b) the corresponding amine (0.5 mmol), TEA (0.5 mmol), dry toluene, 25 °C, 1–24 h.

Linkers with aliphatic amines (morpholine and benzylamine) were synthesized as final products **11** and **12**, respectively. Aromatic amines provided only two linkers bearing *N*-methylaniline and 2-aminopyrimidine (**10** and **13**, respectively). Other heterocyclic amines, such as imidazole, indoline, 2-aminobenzothiazole, 1- and 2-aminobenzoimidazole, and 2-aminobenzoxazole, yielded complex reaction mixtures, as shown in ^{31}P-NMR spectra (Figure S9 in the Supplementary Materials), which we were unable to separate.

Since cargo release was faster in **4–6**, phenyl-lactyl phosphate analogs **14–16** were also prepared with a phenyl instead of an ethyl group attached to the phosphorus.

Then, we subjected **10–16** to irradiation NMR experiments (Figures 5 and 6). Compounds **10–12** afforded the final product **P** within 5 min, and product **P** was a major component in the reaction mixtures in 15 min. In contrast, 2-aminopyrimidine derivative **13** showed a slow formation of intermediate **13-I** without any further spectral change in 15 min. Lastly, the pyrimidine cargo was fully released from **13-I** within 19 days.
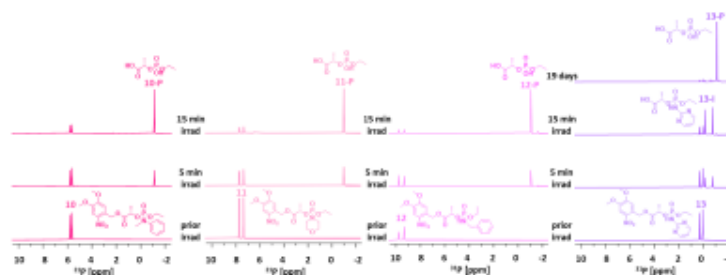
Figure 5. ^{31}P-NMR spectra of linkers **10–13**, measured before and after irradiation with UV light (365 nm) of 5 mM solutions in a solvent mixture of 50% HEPES (pH = 7.4)/DMSO at room temperature (25 °C).

Linkers **14–16** released their amine cargos slightly faster than their ethyl counterparts **7–9**. Linkers **14–16** afforded the final product **P**, which emerged as one singlet ^{31}P signal at $\delta_P \_ 6$ ppm (Figure 6). Although **14** and **15** released the corresponding amines within 5 min, linker **16** did so overnight.
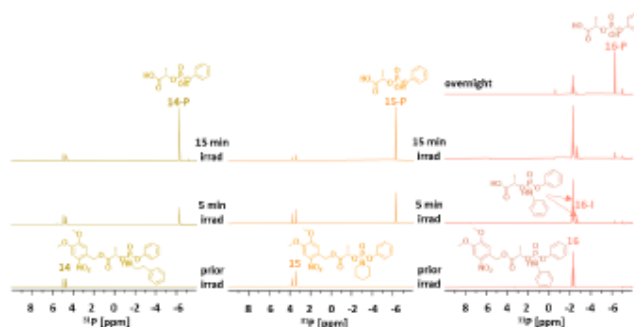
Figure 6. ^{31}P-NMR spectra of linkers **14–16**, measured before and after irradiation with UV light (365 nm) using 5 mM solutions in a solvent mixture of 50% HEPES (pH = 7.4)/DMSO at room temperature (25 °C).

2.5. Application

Ultimately, we prepared two model prodrugs of Ciprofloxacin (Figure 7), which is a known fluoroquinolone antibiotic containing a secondary amino group, to demonstrate the applicability of lactate phosphate-based linkers. To avoid side reactions during the synthesis of **17** and **18**, we protected the carboxylic group of Ciprofloxacin by methylation, which should not decrease the antibiotic activity as reported previously [27]. Then, Ciprofloxacin methyl ester **27** was phosphorylated, following the procedure that had been used for model linkers **7–9** (Scheme 2). Two-step phosphorylation starting from DMNB ester **24** and ethyl dichlorophosphate **20** afforded photoactivable compound **17**. In addition to **17**, we also prepared its enzymatically activable analog **18**. Both compounds were obtained in moderate isolated yields (43–44%) as ca. 1:1 mixtures of diastereoisomers, as confirmed by the presence of two sets of NMR signals.
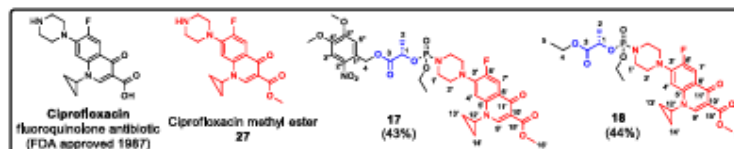
Figure 7. Chemical structures of the FDA-approved fluoroquinolone antibiotic Ciprofloxacin, its methyl ester **27**, and the prodrugs activated by light (**17**) or by an enzyme (**18**) prepared in this work.

The photoactivation of linker **17** resulted in Ciprofloxacin release in 5 min, which was supported by the formation of product **17-P** (Figure 8a). Compound **18** was activated by a lipase from *Candida Antarctica*. We detected **18-P** after 3 h, with ca. 30% cargo release in 24 h. Most Ciprofloxacin (97%) was released in 6 days (Figure 8b). Enzymatic cargo release was relatively slow, which was presumably due to the substrate specificity of the lipase that was used in the experiment.

Furthermore, we performed a biological screening of **18** and **27**, which showed that Ciprofloxacin methylation at the carboxylic moiety inhibits the antibiotic activity of this fluoroquinolone (Table S1 in the Supplementary Materials), despite previous reports stating otherwise [27]. Nevertheless, we believe that phosphate-based linkers may be used to design secondary amine drug delivery systems.
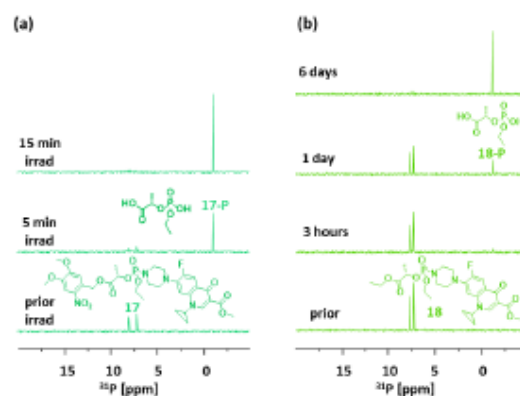
Figure 8. ^{31}P-NMR spectra of Ciprofloxacin prodrugs **17** (a) and **18** (b) recorded in 50% HEPES (pH = 7.4)/DMSO, 25 °C, with linker **17** bearing a photoactivable DMNB group, before and after irradiation by UV light (365 nm), and linker **18**, where SI is activated enzymatically by lipase, before and after lipase addition.

3. Discussion

Our study demonstrates that a universal spacer for delivering all types of amines will unlikely ever be designed given the sensitivity of the phosphorus atom to substitution. An ethylene glycol spacer, which was previously identified as the best linker for phenolic cargo delivery [22], proved inefficient in delivering amines, as shown in **1–6**. An alteration in the electron density of phosphorus caused by oxygen substitution (from phenol—previous work [22]) for nitrogen (from amine) could explain the inefficiency of **1–6**. Although installing phenol (**4–6**) as the second cargo slightly accelerated the SI, when compared to the ethyl analogs **1–3**, phenol was preferentially released instead of the amine cargos. Surprisingly, the L-lactate spacer was the most suitable for aliphatic amines, releasing the cargo in 15 min (**7, 8, 11, 12, 14, 15**). In contrast, linkers with aromatic amine cargos (aniline (**9** and **16**), N-methylaniline (**10**), and 2-aminopyrimidine (**13**)) released their cargos more slowly than linkers bearing aliphatic amines, as indicated by the formation of higher amounts of intermediate **I**.

Slower SI, especially in **13** (**13-I**), could be partly explained by the low pK_a of 2-aminopyrimidine (pK_a 3.54 [28]). Although there is no clear correlation between the pK_a of amine and the SI rate, pK_a plays a key role in amine release [29]. N-protonation in phosphoramidates facilitates the nucleophilic attack of water (a carboxylate group in our case) to the phosphorus atom. Imbach [30] has shown that phosphoramidates consisting of low pK_a amines are stable, whereas those containing amines with a higher pK_a (more than 9) show the fastest hydrolysis, which has also been described by Wagner [31]. Nevertheless, differences in cargo release rate could not be easily explained by the various pK_a of amines. Based on Mayr's extensive studies of amine behavior in solution, the amine leaving group can be affected by attributes other than pK_a, such as nucleophilicity [32], hydration energy (amine stabilization by solvation in water), polarity, and solvent pH, in addition to structural (cyclic vs. acyclic) or sterical effects. Therefore, predicting the optimal spacer for a specific cargo is a difficult task, and designing purposeful drug delivery systems requires studying structure-activity relationships in detail.

4. Materials and Methods

Unless otherwise indicated, all chemicals were purchased from commercial suppliers (Sigma Aldrich, Merck, EU; Fluorochem, UK; Acros Organics, Thermo Fisher Scientific, EU) and used without further purification. All reactions sensitive to air or moisture were

performed under an inert atmosphere of argon in dry solvents. Thin layer chromatography (TLC) was performed on TLC aluminium sheets (silica-gel 60 F_{254}; Merck, EU) and visualized by UV fluorescence. The reaction was monitored by TLC and/or ^{31}P-NMR spectroscopy in $CDCl_3$. Flash-column chromatography was performed on a Compact (ECOM s.r.o., EU) chromatography system using silica-gel or C18 silica-gel 230–400 mesh, 60 Å (Merck KGaA, EU).

NMR spectroscopy. NMR spectra were recorded on a Bruker Avance III spectrometer operating at 400 MHz for ^{1}H and 101 MHz for ^{13}C equipped with a probe with an ATM module (5 mm BBFO BB-19F/1H/D Z-GRD). For NMR signal assignment, standard Bruker pulse sequences were used for 1D (^{1}H, ^{13}C-APT, ^{31}P, ^{19}F) and 2D (COSY, ROESY, HSQC, HMBC) NMR experiments at a corrected temperature of 25 °C. NMR spectra coupled with UV irradiation were recorded on a Bruker Avance III spectrometer with a broad-band cryo probe with an ATM module (5 mm CPBBO BB-1H/19F/15N/D Z-GRD) operating at 500 MHz for ^{1}H and 125.7 MHz for [13]C. All NMR data were interpreted using Topspin 3.5. For reference, the following solvent signals were used: DMSO-d_6: 2.50 (^{1}H) and 39.7 (^{13}C) ppm or $CDCl_3$: 7.28 (^{1}H) and 77.0 (^{13}C) ppm. The ^{31}P-NMR spectra were referenced to H_3PO_4 with 0 ppm.

For NMR experiments with in situ irradiation, a light emitting diode (LED; Thorlabs, EU) was used at 365 nm. The light was guided into the spectrometer, directly into the NMR tube via a multimode silica optical fiber with 1 mm diameter, 0.39 NA, and a high amount of OH (Thorlabs, EU).

Mass spectrometry. Mass spectra were measured on a LTQ. Orbitrap XL (Thermo Fisher Scientific, EU) using electrospray ionization (ESI).

All products were viscous oils, semi-solids or non-crystalline solids. The reaction conditions were not optimized for the highest possible yields.

4,5-dimethoxy-2-nitrobenzyl (2-hydroxyethyl) carbonate (**19**) was prepared according to a literature procedure [22], using 1.0 g of 4,5-dimethoxy-2-nitrobenzyl chloroformate and obtained (0.98 g) as a yellow solid (90%).

L-4,5-dimethoxy-2-nitrobenzyl 2-hydroxypropanoate (**24**) was prepared according to a literature procedure [25], using 2.13 g of 4,5-dimethoxy-2-nitrobenzyl alcohol and obtained (0.94 g) as a light pink solid (33%).

General Procedure for the One-Pot Synthesis of Phosphate-Based Linkers. For each experiment, a DMNB-containing photoarm **19** or **24** (0.5 mmol, 1.0 eq.) was dissolved in 2.5 mL of dry toluene under argon at 25 °C, adding dry TEA (90.6 µL, 0.65 mmol, 1.3 eq.) followed by the corresponding dichlorophosphate (0.5 mmol, 1.0 eq.). The reaction mixture was stirred at 25 °C for 16 h, and the formation of the intermediates was confirmed by ^{31}P-NMR (**22**: δ_P 4.7 ppm; **23**: δ_P 0.08 ppm; **25**: δ_P 4.6 and 4.1 ppm; **26**: δ_P −0.08 and −0.17 ppm) before phosphorylating the amines. The corresponding amine (0.5 mmol, 1.0 eq.) was added, followed by dry TEA (69.7 µL, 0.5 mmol, 1.0 eq.). The reaction mixture was stirred at room temperature until completing the reaction (monitored by ^{31}P-NMR). After evaporating the solvent, pure products were isolated by Flash silica gel chromatography, which was followed by reverse-phase chromatography, as described for each compound.

4,5-dimethoxy-2-nitrobenzyl (2-((ethoxy(phenethylamino)phosphoryl)oxy)ethyl) carbonate (**1**) was prepared from **19** (150.5 mg, 0.5 mmol, 1.0 eq.) and **20** (63.2 µL, 0.5 mmol, 1.0 eq.) according to the general procedure and purified by Flash silica gel chromatography using a gradient (CH_2Cl_2/MeOH 100:0→90:10, v/v), followed by reverse-phase chromatography using a gradient (H_2O/MeCN 100:0→50:50, v/v), which afforded **1** as a dense yellow oil (58.6 mg, 23%). **^{1}H-NMR** (400 MHz, $CDCl_3$, 25 °C) δ 7.74 (s, 1H, 3″), 7.30 (m, 2H, 3′), 7.25–7.20 (m, 1H, 4′), 7.20 (m, 2H, 2′), 7.07 (s, 1H, 6″), 5.60 (s, 1H, 4a), 5.60 (s, 1H, 4b), 4.40 (m, 2H, 2), 4.25–4.10 (m, 2H, 1), 4.10–3.99 (m, 2H, O**CH₂**CH₃), 3.98 and 3.97 (-s, 6H, 4″-O**CH₃**, 5″-O**CH₃**), 3.25–3.16 (m, 2H, NH**CH₂**), 2.80 (m, 1H, **CH₂**Ph), 1.32 (td, 3H, $J_{CH3-CH2}$ = 7.1 Hz, J_{CH3-P} = 0.8 Hz, OCH₂**CH₃**). **^{13}C-NMR** (101 MHz, $CDCl_3$, 25 °C) δ 154.59 (3), 153.52 and 148.38 (4″, 5″), 139.68 (2″), 138.53 (1′), 128.88 (2′), 128.64 (3′), 126.59 (4′), 126.52 (1″), 109.96 (6″), 108.20 (3″), 67.08 (d, J_{2-P} = 7.0 Hz, 2), 66.55 (4), 63.57 (d,

$J_{1\text{-P}}$ = 5.2 Hz, 1), 62.66 (d, $J_{\text{CH2-P}}$ = 5.7 Hz, OCH$_2$CH$_3$), 56.55, and 56.44 (4″-OCH$_3$, 5″-OCH$_3$), 42.65 (NHCH$_2$), 37.85 (d, $J_{\text{CH2-P}}$ = 6.2 Hz, **CH$_2$**Ph), 16.19 (d, $J_{\text{CH3-P}}$ = 6.9 Hz, OCH$_2$**CH$_3$**). ^{31}**P-NMR** (162 MHz, CDCl$_3$, 25 °C) δ 9.1 ppm. **HR-MS** (ESI) calculated for C$_{22}$H$_{30}$O$_{10}$N$_2$P 513.16326, found [M + H]$^+$ 513.16284.

4,5-dimethoxy-2-nitrobenzyl (2-((ethoxy(piperidin-1-yl)phosphoryl)oxy)ethyl) carbonate (**2**) was prepared from **19** (150.5 mg, 0.5 mmol, 1.0 eq.) and **20** (63.2 µL, 0.5 mmol, 1.0 eq.) according to the general procedure and purified by Flash silica gel chromatography using a gradient (CH$_2$Cl$_2$/MeOH 100:0→90:10, *v*/*v*), followed by reverse-phase chromatography using a gradient (H$_2$O/MeCN 100:0→50:50, *v*/*v*), which afforded **2** as a dense yellow oil (103.7 mg, 0.22 mmol, 44%). ^{1}**H-NMR** (400 MHz, CDCl$_3$, 25 °C) δ 7.74 (s, 1H, 3″), 7.09 (s, 1H, 6″), 5.62–5.60 (m, 2H, 4), 4.46–4.35 (m, 2H, 2), 4.28–4.12 (m, 2H, 1), 4.09–3.98 (m, 2H, OCH$_2$CH$_3$), 4.00 (s, 3H, 5″-OCH$_3$), 3.97 (s, 3H, 4″-OCH$_3$), 3.14–3.05 (m, 4H, 1′), 1.62–1.47 (m, 6H, 2′, 3′), 1.31 (td, $J_{\text{CH3-CH2}}$ = 7.1 Hz, $J_{\text{CH3-P}}$ = 0.8 Hz, OCH$_2$CH$_3$). ^{13}**C-NMR** (101 MHz, CDCl$_3$, 25 °C) δ 154.58 (3), 153.73 (5″), 148.35 (4″), 139.65 (2″), 126.62 (1″), 109.89 (6″), 108.18 (3″), 67.14 (d, $J_{\text{2-P}}$ = 7.4 Hz, 2), 66.47 (4), 63.38 (d, $J_{1\text{-P}}$ = 5.2 Hz, 1), 62.39 (d, $J_{\text{CH2-P}}$ = 5.9 Hz, OCH$_2$CH$_3$), 56.55 (5″-OCH$_3$), 56.43 (4″-OCH$_3$), 45.36 (d, $J_{1'\text{-P}}$ = 2.3 Hz, 1′), 26.02 (d, $J_{2'\text{-P}}$ = 5.1 Hz, 2′), 24.39 (d, $J_{3'\text{-P}}$ = 1.5 Hz, 3′), 16.17 (d, $J_{\text{CH3-CH2}}$ = 7.0 Hz, OCH$_2$**CH$_3$**). ^{31}**P-NMR** (162 MHz, CDCl$_3$, 25 °C) δ 9.10 ppm. **HR-MS** (ESI) calculated for C$_{19}$H$_{30}$O$_{10}$N$_2$P 477.16326, found [M + H]$^+$ 477.16340.

4,5-dimethoxy-2-nitrobenzyl (2-((ethoxy(phenylamino)phosphoryl)oxy)ethyl) carbonate (**3**) was prepared from **19** (150.5 mg, 0.5 mmol, 1.0 eq.) and **20** (63.2 µL, 0.5 mmol, 1.0 eq.) according to the general procedure and purified by Flash silica gel chromatography using a gradient (CH$_2$Cl$_2$/MeOH 100:0→90:10, *v*/*v*), followed by two reverse-phase chromatographies using a gradient (H$_2$O/MeCN 100:0→50:50, *v*/*v*), which afforded **3** as a dense yellow oil (82.8 mg, 34%). ^{1}**H-NMR** (400 MHz, CDCl$_3$, 25 °C) δ 7.75 (s, 1H, 3″), 7.24 (m, 2H, 3′), 7.06 (s, 1H, 6″), 7.00 (m, 2H, 2′), 6.95 (m, 1H, 4′), 5.75 (d, $J_{\text{NH-P}}$ = 9.0 Hz, NH), 5.56 (m, 2H, 4), 4.42 (m, 2H, 2), 4.41–4.25 (m, 2H, 1), 4.25–4.08 (m, 2H, OCH$_2$CH$_3$), 3.98–3.97 (m, 6H, 4″-OCH$_3$, 5″-OCH$_3$), 1.33 (td, $J_{\text{CH3-CH2}}$ = 7.1 Hz, $J_{\text{CH3-P}}$ = 0.8 Hz, OCH$_2$CH$_3$). ^{13}**C-NMR** (101 MHz, CDCl$_3$, 25 °C) δ 154.49 (3), 153.76 and 148.34 (4″, 5″), 139.61 (2″), 139.16 (1′), 129.35 (3′), 126.58 (1″), 121.97 (4′), 117.45 (d, $J_{2'\text{-P}}$ = 7.2 Hz, 2′), 109.82 (6″), 108.18 (3″), 66.78 (d, $J_{\text{2-P}}$ = 7.3 Hz, 2), 66.55 (4), 64.00 (d, $J_{1\text{-P}}$ = 4.6 Hz, 1), 63.31 (d, $J_{\text{CH2-P}}$ = 5.0 Hz, OCH$_2$CH$_3$), 56.57 and 56.43 (4″-OCH$_3$, 5″-OCH$_3$), 16.07 (d, $J_{\text{CH3-P}}$ = 7.0 Hz, OCH$_2$**CH$_3$**). ^{31}**P-NMR** (162 MHz, CDCl$_3$, 25 °C) δ 2.31 ppm. **HR-MS** (ESI) calculated for C$_{20}$H$_{26}$O$_{10}$N$_2$P 485.13196, found [M + H]$^+$ 485.13189.

4,5-dimethoxy-2-nitrobenzyl (2-(((phenethylamino)(phenoxy)phosphoryl)oxy)ethyl) carbonate (**4**) was prepared from **19** (150.5 mg, 0.5 mmol, 1.0 eq.) and **21** (74.7 µL, 0.5 mmol, 1.0 eq.) according to the general procedure and purified by Flash silica gel chromatography using a gradient (hexane/EtOAc 100:0→0:100, *v*/*v*), followed by reverse-phase chromatography using a gradient (H$_2$O/MeCN 100:0→0:100, *v*/*v*) and preparative HPLC chromatography using a gradient (H$_2$O/MeCN 95:5 -> 20:80, *v*/*v*), which afforded **4** as a dense yellow oil (10.2 mg, 4%). ^{1}**H-NMR** (400 MHz, CDCl$_3$, 25 °C) δ 7.74 (s, 1H, 3″), 7.35–7.12 (m, 10H, 2′, 3′, 4′, 2‴, 3‴, 4‴), 7.06 (s, 1H, 6″), 5.60 (s, 1H, 4a), 5.60 (s, 1H, 4b), 4.47–4.37 (m, 2H, 2), 4.35–4.21 (m, 2H, 1), 3.97 (s, 3H, 4″-OCH$_3$), 3.94 (s, 3H, 5″-OCH$_3$), 3.33–3.24 (m, 2H, NHCH$_2$), 2.93–2.82 (m, 1H, NH), 2.77 (t, 2H, $J_{\text{CH2Ph-CH2NH}}$ = 6.8 Hz, **CH$_2$**Ph). ^{13}**C-NMR** (101 MHz, CDCl$_3$, 25 °C) δ 154.55 (3), 153.73 (5″), 150.78 (d, $J_{1'''\text{-P}}$ = 6.8 Hz, 1‴), 148.37 (4″), 139.65 (2″), 138.27 (1′), 129.69 (3′ or 3‴), 128.87 (2′), 128.67 (3′ or 3‴), 126.65 (4′), 126.50 (1″), 124.87 (4‴), 120.13 (d, $J_{2'''\text{-P}}$ = 4.8 Hz, 2‴), 109.94 (6″), 108.19 (3″), 66.87 (d, $J_{\text{2-P}}$ = 7.4 Hz, 2), 66.58 (4), 64.18 (d, $J_{1\text{-P}}$ = 5.2 Hz, 1), 56.54 (5″-OCH$_3$), 56.43 (4″-OCH$_3$), 42.78 (NHCH$_2$), 37.69 (d, $J_{\text{CH2-P}}$ = 6.0 Hz, **CH$_2$**Ph). ^{31}**P-NMR** (162 MHz, CDCl$_3$, 25 °C) δ 4.42 ppm. **HR-MS** (ESI) calculated for C$_{26}$H$_{30}$O$_{10}$N$_2$P 561.16326, found [M + H]$^+$ 561.16318.

4,5-dimethoxy-2-nitrobenzyl (2-(((piperidin-1-yl)(phenoxy)phosphoryl)oxy)ethyl) carbonate (**5**) was prepared from **19** (150.5 mg, 0.5 mmol, 1.0 eq.) and **21** (74.7 µL, 0.5 mmol, 1.0 eq.) according to the general procedure and purified by Flash silica gel chromatography using a gradient (hexane/EtOAc 100:0→0:100, *v*/*v*), followed by reverse-phase chromatography

using a gradient (H$_2$O/MeCN 100:0→0:100, *v*/*v*) and preparative HPLC chromatography using a gradient (H$_2$O/MeCN 95:5→0:100, *v*/*v*), which afforded **5** as a dense yellow oil (31.9 mg, 12%). **^{1}H-NMR** (400 MHz, CDCl$_3$, 25 °C) δ 7.75 (s, 1H, 3″), 7.32 (m, 2H, 3‴), 7.22 (m, 2H, 2‴), 7.14 (m, 1H, 4‴), 7.09 (s, 1H, 6″), 5.62 (s, 1H, 4a), 5.62 (s, 1H, 4b), 4.47–4.43 (m, 2H, 2), 4.35–4.28 (m, 2H, 1), 3.99 (s, 3H, 5″-OCH$_3$), 3.98 (s, 3H, 4″-OCH$_3$), 3.20–3.12 (m, 4H, 1′), 1.60–1.51 (m, 2H, 3′), 1.51–1.44 (m, 4H, 2′). **^{13}C-NMR** (101 MHz, CDCl$_3$, 25 °C) δ 154.57 (3), 153.75 (5″), 150.98 (d, $J_{1′′′\text{-}P}$ = 6.8 Hz, 1‴), 148.36 (4″), 139.64 (2″), 129.60 (3‴), 126.61 (1″), 124.62 (4″), 120.07 (d, $J_{2′′′\text{-}P}$ = 5.2 Hz, 2‴), 109.90 (6″), 108.18 (3″), 66.96 (d, $J_{2\text{-}P}$ = 7.3 Hz, 2), 66.53 (4), 63.98 (d, $J_{1\text{-}P}$ = 5.3 Hz, 1), 56.55 and 56.44 (4″-OCH$_3$, 5″-OCH$_3$), 45.49(d, $J_{1′\text{-}P}$ = 2.3 Hz, 1′), 25.82 (d, $J_{2′\text{-}P}$ = 4.7 Hz, 2′), 25.26 (3′). **^{31}P-NMR** (162 MHz, CDCl$_3$, 25 °C) δ 4.27 ppm. **HR-MS** (ESI) calculated for C$_{23}$H$_{30}$O$_{10}$N$_2$P 525.16326, found [M + H]$^+$ 525.16278.

4,5-dimethoxy-2-nitrobenzyl 2-(((phenylamino)(phenoxy)phosphoryl)oxy)ethyl carbonate (**6**) was prepared from **19** (150.5 mg, 0.5 mmol, 1.0 eq.) and **21** (74.7 μL, 0.5 mmol, 1.0 eq.) according to the general procedure and purified by Flash silica gel chromatography using a gradient (hexane/EtOAc 100:0→0:100, *v*/*v*), followed by reverse-phase chromatography using a gradient (H$_2$O/MeCN 100:0→0:100, *v*/*v*) and preparative HPLC chromatography using a gradient (H$_2$O/MeCN 95:5→0:100, *v*/*v*), which afforded **6** as a dense yellow oil (18.1 mg, 7%). **^{1}H-NMR** (400 MHz, CDCl$_3$, 25 °C) δ 7.74 (s, 1H, 3″), 7.28–7.22 (m, 4H, 3′, 3‴), 7.17–7.11 (m, 3H, 2‴, 4‴), 7.07 (m, 2H, 2′), 7.05 (s, 1H, 6″), 6.99 (m, 1H, 4′), 6.25 (d, $J_{\text{NH-P}}$ = 9.8 Hz, NH), 5.62–5.52 (m, 2H, 4), 4.49–4.33 (m, 4H, 1, 2), 3.97 (s, 3H, 4″-OCH$_3$), 3.93 (s, 3H, 5″-OCH$_3$). **^{13}C-NMR** (101 MHz, CDCl$_3$, 25 °C) δ 154.47 (3), 153.77 (5″), 150.15 (d, $J_{1′′′\text{-}P}$ = 6.1 Hz, 1‴), 148.32 (4″), 139.56 (2″), 138.79 (1′), 129.70 (3‴), 129.39 (3′), 126.58 (1″), 125.30 (4‴), 122.33 (4′), 120.38 (d, $J_{2′′′\text{-}P}$ = 4.8 Hz, 2‴), 117.87 (d, $J_{2′\text{-}P}$ = 7.4 Hz, 2′), 109.77 (6″), 108.16 (3″), 66.61 (d, $J_{2\text{-}P}$ = 7.0 Hz, 2), 66.59 (4), 64.53 (d, $J_{1\text{-}P}$ = 4.6 Hz, 1), 56.53 (5″-OCH$_3$), 56.42 (4″-OCH$_3$). **^{31}P-NMR** (162 MHz, CDCl$_3$, 25 °C) δ −1.97 ppm. **HR-MS** (ESI) calculated for C$_{24}$H$_{26}$O$_{10}$N$_2$P 533.13196, found [M + H]$^+$ 533.13188.

4,5-dimethoxy-2-nitrobenzyl (2S)-2-((ethoxy(phenethylamino)phosphoryl)oxy)propanoate (**7**) was prepared from **24** (142.6 mg, 0.5 mmol, 1.0 eq.) and **20** (63.2 μL, 0.5 mmol, 1.0 eq.) according to the general procedure and purified by Flash silica gel chromatography using a gradient (hexane/EtOAc 50:50→0:100, *v*/*v*), followed by reverse-phase chromatography using a gradient (H$_2$O/MeOH 100:0→0:100, *v*/*v*), which afforded **7** as a dense yellow oil (129.3 mg, 52%). **^{1}H-NMR** (400 MHz, CDCl$_3$, 25 °C, 2 diastereoisomers in ca. 1:1 ratio) δ 7.73 and 7.72 (s, 2H, 3″), 7.33–7.13 (m, 10H, 2′, 3′, 4′), 7.07 and 7.03 (s, 2H, 6″), 5.66–5.53 (m, 4H, 4), 5.02–4.88 (m, 2H, 1), 4.14–3.94 (m, 4H, OCH$_2$CH$_3$), 4.00 and 3.98 (s, 6H, 5″-OCH$_3$), 3.96 and 3.96 (s, 6H, 4″-OCH$_3$), 3.26–3.16 (m, 4H, NHCH$_2$), 2.83–2.67 (m, 6H, CH$_2$Ph, NH), 1.61 (d, 3H, $J_{2\text{-}1}$ = 6.8 Hz), 1.56 (d, 3H, $J_{2\text{-}1}$ = 6.9 Hz, 2), 1.32 (td, 3H, $J_{\text{CH3-CH2}}$ = 7.0 Hz, $J_{\text{CH3-P}}$ = 0.8 Hz, OCH$_2$CH$_3$), 1.28 (td, 3H, $J_{\text{CH3-CH2}}$ = 7.2 Hz, $J_{\text{CH3-P}}$ = 0.7 Hz, OCH$_2$CH$_3$). **^{13}C-NMR** (101 MHz, CDCl$_3$, 25 °C) δ 170.68–170.47 (m, 3), 153.65 and 153.63 (5″), 148.40 and 148.37 (4″), 139.84 (2″), 138.55 and 138.50 (1′), 128.86 and 128.83 (2′), 128.65 and 128.58 (3′), 126.53 and 126.43 (4′), 126.32 and 126.29 (1″), 110.47 and 110.35 (6″), 108.26 and 108.23 (3″), 70.57 (d, $J_{1\text{-}P}$ = 7.6 Hz, 1), 70.52 (d, $J_{1\text{-}P}$ = 7.6 Hz, 1), 64.09 and 64.06 (4), 62.93 (d, $J_{\text{CH2-}P}$ = 5.7 Hz, OCH$_2$CH$_3$), 62.79 (d, $J_{\text{CH2-}P}$ = 5.7 Hz, OCH$_2$CH$_3$), 56.68 and 56.62 (4″-OCH$_3$), 56.42 (5″-OCH$_3$), 42.65 and 42.54 (NHCH$_2$), 37.88–37.70 (m, CH$_2$Ph), 19.42 (d, $J_{2\text{-}P}$ = 5.5 Hz, 2), 19.41 (d, $J_{2\text{-}P}$ = 5.3 Hz, 2), 16.18 (d, $J_{\text{CH3-}P}$ = 7.0 Hz, OCH$_2$CH$_3$), 16.13 (d, $J_{\text{CH3-}P}$ = 7.0 Hz, OCH$_2$CH$_3$). **^{31}P-NMR** (162 MHz, CDCl$_3$, 25 °C) δ 9.07 and 8.39 ppm. **HR-MS** (ESI) calculated for C$_{22}$H$_{29}$O$_9$N$_2$NaP 519.15029, found [M + Na]$^+$ 519.14986.

4,5-dimethoxy-2-nitrobenzyl (2S)-2-((ethoxy(piperidin-1-yl)phosphoryl)oxy)propanoate (**8**) was prepared from **24** (142.6 mg, 0.5 mmol, 1.0 eq.) and **20** (63.2 μL, 0.5 mmol, 1.0 eq.) according to the general procedure and purified by Flash silica gel chromatography using a gradient (CH$_2$Cl$_2$/MeOH 100:0→80:20, *v*/*v*), followed by reverse-phase chromatography using a gradient (H$_2$O/MeCN 100:0→0:100, *v*/*v*), which afforded **8** as a dense yellow oil (97.0 mg, 42%). **^{1}H-NMR** (400 MHz, CDCl$_3$, 25 °C, 2 diastereoisomers in ca. 1:1 ratio) δ 7.74 and 7.73 (s, 2H, 3″), 7.10 and 7.09 (s, 2H, 6″), 5.64 (dd, 1H, J_{Gem} = 14.8 Hz, $J_{4a\text{-}6′′}$ = 0.6 Hz,

4a), 5.63 (dd, 1H, J_{Gem} = 14.8 Hz, $J_{4a\text{-}6''}$ = 0.6 Hz, 4a), 5.59 (dd, 1H, J_{Gem} = 14.8 Hz, $J_{4b\text{-}6''}$ = 0.6 Hz, 4b), 5.58 (dd, 1H, J_{Gem} = 14.8 Hz, $J_{4b\text{-}6''}$ = 0.6 Hz, 4b), 5.00–4.88 (m, 1H, 1), 4.12–3.99 (m, 2H, O**CH₂**CH₃), 4.02 and 4.02 (s, 6H, 5''-O**CH₃**), 3.9 and 3.97 (s, 6H, 4''-O**CH₃**), 3.19–3.01 (m, 4H, 1'), 1.61 (d, $J_{2\text{-}P}$ = 6.9 Hz, 3H, 2), 1.59 (d, $J_{2\text{-}P}$ = 6.9 Hz, 3H, 2), 1.57–1.44 (m, 10H, 2', 3'), 1.33 (td, 3H, $J_{CH3\text{-}CH2}$ = 7.1 Hz, $J_{CH3\text{-}P}$ = 0.8 Hz, O**CH₂CH₃**), 1.28 (td, 3H, $J_{CH3\text{-}CH2}$ = 7.1 Hz, $J_{CH3\text{-}P}$ = 0.8 Hz, O**CH₂CH₃**). **¹³C-NMR** (101 MHz, CDCl₃, 25 °C) δ 170.71 (d, $J_{3\text{-}P}$ = 4.4 Hz, 3), 170.51 (d, $J_{3\text{-}P}$ = 5.7 Hz, 3), 153.70 and 153.68 (5''), 148.28 (4''), 139.74 (2''), 126.66 and 126.60 (1''), 110.37 and 110.24 (3''), 70.33 (d, $J_{1\text{-}P}$ = 5.1 Hz, 1), 70.18 (d, $J_{1\text{-}P}$ = 5.2 Hz, 1), 63.92 and 63.89 (4), 62.60 (d, $J_{CH2\text{-}P}$ = 5.9 Hz, O**CH₂**CH₃), 62.42 (d, $J_{CH2\text{-}P}$ = 5.8 Hz, O**CH₂**CH₃), 56.66 and 56.61 and 56.39–56.36 (m) (4''-O**CH₃**, 5''-O**CH₃**), 45.34 and 45.32 (1'), 25.96 (d, $J_{2'\text{-}P}$ = 4.7 Hz, 2'), 25.84 (d, $J_{2'\text{-}P}$ = 5.4 Hz, 2'), 24.33 (d, $J_{3'\text{-}P}$ = 1.5 Hz, 3'), 24.29 (d, $J_{3'\text{-}P}$ = 1.5 Hz, 3'), 19.45 (d, $J_{2\text{-}P}$ = 4.8 Hz, 2), 19.41 (d, $J_{2\text{-}P}$ = 5.2 Hz, 2), 16.17–16.01 (m, O**CH₂CH₃**). **³¹P-NMR** (162 MHz, CDCl₃, 25 °C) δ 8.81 and 8.30 ppm. **HR-MS** (ESI) calculated for C₁₉H₂₉O₉N₂NaP 483.15029, found [M + Na]⁺ 483.14977.

4,5-dimethoxy-2-nitrobenzyl (2S)-2-((ethoxy(phenylamino)phosphoryl)oxy)propanoate **(9)** was prepared from **24** (142.6 mg, 0.5 mmol, 1.0 eq.) and **20** (63.2 μL, 0.5 mmol, 1.0 eq.) according to the general procedure and purified by Flash silica gel chromatography using a gradient (hexane/EtOAc 100:0→0:100, v/v), followed by reverse-phase chromatography using a gradient (H₂O/MeCN 100:0→0:100, v/v), which afforded **9** as a dense yellow oil (108.2 mg, 46%). **¹H-NMR** (400 MHz, CDCl₃, 25 °C, 2 diastereoisomers in ca. 1:1 ratio) δ 7.73 and 7.64 (s, H, 3''), 7.22 and 7.17 (m, 4H, 3'), 7.08 and 6.94 (s, 2H, 6''), 7.06–6.96 (m, 4H, 2'), 6.96–6.89 (m, 4H, 4'), 6.35 (d, 1H, $J_{NH\text{-}P}$ = 9.3 Hz, NH), 6.22 (d, 1H, $J_{NH\text{-}P}$ = 8.7 Hz, NH), 5.65 (dd, 1H, J_{Gem} = 14.9 Hz, $J_{4a\text{-}6''}$ = 0.6 Hz, 4a), 5.57 (dd, 1H, J_{Gem} = 14.9 Hz, $J_{4b\text{-}6''}$ = 0.5 Hz, 4b), 5.56–5.41 (m, 2H, 4a, 4b), 5.12 and 5.06 (m, 2H, 1), 4.29–4.07 (m, 4H, O**CH₂**CH₃), 3.97–3.95 (m, 6H, 4''-O-**CH₃**), 3.95 and 3.90 (s, 6H, 5''-O**CH₃**), 1.67 (d, 3H, $J_{2\text{-}1}$ = 6.9 Hz, 2), 1.52 (d, 3H, $J_{2\text{-}1}$ = 6.9 Hz, 2), 1.36–1.26 (m, 6H, O**CH₂CH₃**). **¹³C-NMR** (101 MHz, CDCl₃, 25 °C) δ 170.13 (d, $J_{3\text{-}P}$ = 4.6 Hz, 3), 170.10 (d, $J_{3\text{-}P}$ = 5.4 Hz, 3), 153.64 and 153.55 (5''), 148.28 and 148.26 (4''), 139.69 and 139.67 (2''), 139.20 (d, $J_{1'\text{-}P}$ = 4.7 Hz, 1'), 129.19 and 129.10 (3'), 126.40 and 126.16 (1''), 121.88 and 121.84 (4'), 117.63 (d, $J_{2'\text{-}P}$ = 7.7 Hz, 2'), 117.55 (d, $J_{2'\text{-}P}$ = 7.3 Hz, 2'), 110.30 and 110.17 (6''), 108.15 and 108.11 (3''), 71.25 (d, $J_{1\text{-}P}$ = 4.6 Hz, 1), 70.92 (d, $J_{1\text{-}P}$ = 4.7 Hz, 1), 64.12 and 64.09 (4), 63.46 (d, $J_{CH2\text{-}P}$ = 5.4 Hz, O**CH₂CH₃**), 63.36 (d, $J_{CH2\text{-}P}$ = 5.6 Hz, O**CH₂CH₃**), 56.53, 56.40, 56.35, and 56.33 (4''-O**CH₃**, 5''-O**CH₃**), 19.30 (d, $J_{2\text{-}P}$ = 5.1 Hz, 2), 19.12 (d, $J_{2\text{-}P}$ = 6.0 Hz, 2), 15.98 (d, $J_{CH3\text{-}P}$ = 7.3 Hz, O**CH₂CH₃**), 15.95 (d, $J_{CH3\text{-}P}$ = 7.3 Hz, O**CH₂CH₃**). **³¹P-NMR** (162 MHz, CDCl₃, 25 °C) δ 2.14 and 1.86 ppm. **HR-MS** (ESI) calculated for C₂₀H₂₅O₉N₂NaP 491.11899, found [M + Na]⁺ 491.11862.

4,5-dimethoxy-2-nitrobenzyl (2S)-2-((ethoxy(N-methylphenyl)phosphoryl)oxy)propanoate **(10)** was prepared from **24** (142.6 mg, 0.5 mmol, 1.0 eq.) and **20** (63.2 μL, 0.5 mmol, 1.0 eq.) according to the general procedure and purified by Flash silica gel chromatography using a gradient (hexane/EtOAc 100:0→0:100, v/v), followed by two reverse-phase chromatographies using a gradient (H₂O/MeCN 100:0 -> 50:50, v/v), which afforded **10** as a dense yellow oil (44.5 mg, 18%). **¹H-NMR** (400 MHz, CDCl₃, 25 °C, 2 diastereoisomers in ca. 1:1 ratio) δ 7.74 and 7.73 (s, 2H, 3''), 7.33–7.30 (m, 2H, 2'), 7.30–7.22 (m, 6H, 2', 3'), 7.14–7.04 (m, 2H, 4'), 7.10 and 7.00 (s, 2H, 6''), 5.65–5.62 and 5.56–5.54 (m, 4H, 4), 5.11–4.96 (m, 2H, 1), 4.23–4.00 (m, 4H, O**CH₂**CH₃), 3.99 and 3.98 3.98 (s, 9H, 4''-O**CH₃**, 5''-O**CH₃**), 3.91 (s, 3H, 5''-O**CH₃**), 3.22 (d, 3H, $J_{CH3\text{-}P}$ = 9.0 Hz, N**CH₃**), 3.17 (d, 3H, $J_{CH3\text{-}P}$ = 9.1 Hz, N**CH₃**), 1.65 (d, 3H, $J_{2\text{-}1}$ = 7.0 Hz, 2), 1.51 (d, 3H, $J_{2\text{-}1}$ = 7.0 Hz, 2), 1.30 (td, 3H, $J_{CH3\text{-}CH2}$ = 7.0 Hz, $J_{CH3\text{-}P}$ = 0.8 Hz, O**CH₂CH₃**), 1.28 (dt, 3H, $J_{CH3\text{-}CH2}$ = 7.0 Hz, $J_{CH3\text{-}P}$ = 0.8 Hz, O**CH₂CH₃**). **¹³C-NMR** (101 MHz, CDCl₃, 25 °C) δ 170.4* and 170.2* (3), 153.8* (5''), 148.3* (4''), 143.9* (1'), 139.8* and 139.7* (2''), 128.97 and 128.85 (3'), 126.67 and 126.65 (1''), 124.27 and 124.19 (4'), 123.13 (d, $J_{2'\text{-}P}$ = 4.6 Hz, 2'), 122.69 (d, $J_{2'\text{-}P}$ = 3.9 Hz, 2'), 110.33 and 110.11 (6''), 108.21 and 108.15 (3''), 71.15 (d, $J_{1\text{-}P}$ = 5.4 Hz, 1), 70.90 (d, $J_{1\text{-}P}$ = 4.9 Hz, 1), 64.08 and 64.04 (4), 63.34 (d, $J_{CH2\text{-}P}$ = 5.7 Hz, O**CH₂CH₃**), 62.96 (d, $J_{CH2\text{-}P}$ = 5.6 Hz, O**CH₂CH₃**), 56.69 and 56.58 and 56.49–56.33 (m, 4''-O**CH₃**, 5''-O**CH₃**), 37.35–36.95 (m, N**CH₃**), 19.37 (d, $J_{2\text{-}P}$ = 4.6 Hz, 2),

19.16 (d, $J_{2\text{-}P}$ = 5.4 Hz, 2), 15.99 (d, $J_{\text{CH3-P}}$ = 6.7 Hz, OCH$_2$**CH$_3$**). * The ^{13}C chemical shift was extracted from HMBC. ^{31}**P-NMR** (162 MHz, CDCl$_3$, 25 °C) δ 5.82 and 5.55 ppm. **HR-MS** (ESI) calculated for C$_{21}$H$_{27}$O$_9$N$_2$NaP 505.13464, found [M + Na]$^+$ 505.13472.

4,5-dimethoxy-2-nitrobenzyl (2S)-2-((ethoxy(morpholin-1-yl)phosphoryl)oxy)propanoate **(11)** was prepared from **24** (142.6 mg, 0.5 mmol, 1.0 eq.) and **20** (63.2 µL, 0.5 mmol, 1.0 eq.) according to the general procedure and purified by Flash silica gel chromatography using a gradient (hexane/EtOAc 100:0→0:100, v/v), followed by two reverse-phase chromatographies using a gradient (H$_2$O/MeCN 100:0→60:40, v/v), which afforded **11** as a dense yellow oil (80.0 mg, 35%). ^{1}**H-NMR** (400 MHz, CDCl$_3$, 25 °C, 2 diastereoisomers in ca. 1:1 ratio) δ 7.72–7.70 (m, 2H, 3″), 7.04 and 7.02 (s, 2H, 6″), 5.64–5.51 (m, 4H, 4), 5.00–4.87 (m, 2H, 1), 4.13–4.01 (m, 4H, OCH$_2$CH$_3$), 3.99 and 3.99 (s, 6H, 5″-OCH$_3$), 3.96–3.94 (m, 6H, 4″-OCH$_3$), 3.64 and 3.59 (m, 8H, 2′), 3.19–3.04 (m, 8H, 1′), 1.59 (dd, 6H, $J_{2\text{-}1}$ = 6.9 Hz, $J_{2\text{-}P}$ = 1.0 Hz, 2), 1.32 (td, 3H, $J_{\text{CH3-CH2}}$ = 7.1 Hz, $J_{\text{CH3-P}}$ = 0.9 Hz, OCH$_2$**CH$_3$**), 1.28 (td, 3H, $J_{\text{CH3-CH2}}$ = 7.1 Hz, $J_{\text{CH3-P}}$ = 0.8 Hz, OCH$_2$**CH$_3$**). ^{13}**C-NMR** (101 MHz, CDCl$_3$, 25 °C) δ 170.41 (d, $J_{3\text{-}P}$ = 4.6 Hz, 3), 170.27 (d, $J_{3\text{-}P}$ = 4.9 Hz, 3), 153.55 and 153.52 (5″), 148.37 and 148.33 (4″), 139.85 and 139.82 (2″), 126.21 and 126.08 (1″), 110.53 and 110.49 (6″), 108.20 and 108.15 (3″), 70.49 (d, $J_{1\text{-}P}$ = 5.0 Hz, 1), 70.40 (d, $J_{1\text{-}P}$ = 5.3 Hz, 1), 66.83 (d, $J_{2'\text{-}P}$ = 5.8 Hz, 2′), 66.73 (d, $J_{2'\text{-}P}$ = 6.1 Hz, 2′), 64.02 and 64.00 (4), 62.95 (d, $J_{\text{CH2-P}}$ = 6.1 Hz, OCH$_2$**CH$_3$**), 62.74 (d, $J_{\text{CH2-P}}$ = 6.1 Hz, OCH$_2$**CH$_3$**), 56.56 and 56.52 (5″-OCH$_3$), 56.35–56.30 (m, 4″-OCH$_3$), 44.49–44.43 (m, 1′), 19.34 (d, $J_{2\text{-}P}$ = 4.7 Hz, 2), 19.33 (d, $J_{2\text{-}P}$ = 5.5 Hz, 2), 16.03 (d, $J_{\text{CH3-P}}$ = 7.0 Hz, OCH$_2$**CH$_3$**), 15.98 (d, $J_{\text{CH3-P}}$ = 6.7 Hz, OCH$_2$**CH$_3$**). ^{31}**P-NMR** (162 MHz, CDCl$_3$, 25 °C) δ 7.68 and 7.13 ppm. **HR-MS** (ESI) calculated for C$_{18}$H$_{28}$O$_{10}$N$_2$P 463.14761, found [M + H]$^+$ 463.14720.

4,5-dimethoxy-2-nitrobenzyl (2S)-2-((ethoxy(benzylamino)phosphoryl)oxy)propanoate **(12)** was prepared from **24** (142.6 mg, 0.5 mmol, 1.0 eq.) and **20** (63.2 µL, 0.5 mmol, 1.0 eq.) according to the general procedure and purified by Flash silica gel chromatography using a gradient (hexane/EtOAc 50:50→0:100, v/v), followed by reverse-phase chromatography using a gradient (H$_2$O/MeOH 100:0→0:100, v/v), which afforded **12** as a dense yellow oil (101 mg, 42%). ^{1}**H-NMR** (400 MHz, CDCl$_3$, 25 °C, 2 diastereoisomers in ca. 1:1 ratio) δ 7.73 and 7.71 (s, 2H, 3″), 7.36–7.22 (m, 10H, 2′, 3′, 4′), 7.07 and 7.02 (s, 2H, 6″), 5.63 (d, 1H, J_{Gem} = 14.6 Hz, 4a), 5.58 (d, J_{Gem} = 14.6 Hz, 1H, 4b), 5.56 (d, 1H, J_{Gem} = 14.7 Hz, 4a), 5.51 (d, 1H, J_{Gem} = 14.7 Hz, 4b), 5.04 and 4.97 (m, 2H, 1), 4.17–4.05 (m, 8H, NHCH$_2$, OCH$_2$CH$_3$), 4.00 and 3.98 (s, 6H, 5″-OCH$_3$), 3.96 and 3.96 (s, 6H, 4″-OCH$_3$), 3.12 (m, 2H, NH), 1.63 (d, 3H, $J_{2\text{-}1}$ = 6.9 Hz, 2), 1.54 (d, 3H, $J_{2\text{-}1}$ = 6.9 Hz, 2), 1.32 (tm, $J_{\text{CH3-CH2}}$= 7.1 Hz, 3H, OCH$_2$**CH$_3$**), 1.29 (tm, $J_{\text{CH3-CH2}}$= 7.1 Hz, 3H, OCH$_2$**CH$_3$**). ^{13}**C-NMR** (101 MHz, CDCl$_3$, 25 °C) δ 170.69–170.57 (m, 3), 153.66 and 153.61 (5″), 148.35 and 148.32 (4″), 139.80 (2″), 139.42–139.30 (m, 1′), 128.56 and 128.48 (3′), 127.41, 127.30, and 127.21 (2′, 4′), 126.43 and 126.23 (1″), 110.46 and 110.37 (6″), 108.19 (3″), 70.67 (d, $J_{1\text{-}P}$ = 4.7 Hz, 1), 70.60 (d, $J_{1\text{-}P}$ = 5.1 Hz, 1), 64.07 and 64.04 (4), 63.06 (d, $J_{\text{CH2-P}}$ = 5.4 Hz, OCH$_2$**CH$_3$**), 62.91 (d, $J_{\text{CH2-P}}$ = 5.4 Hz, OCH$_2$**CH$_3$**), 56.63 and 56.55 (5″-OCH$_3$), 56.37 (4″-OCH$_3$), 45.23 (d, $J_{\text{CH2-P}}$ = 6.2 Hz, NHCH$_2$), 19.34 (d, $J_{2\text{-}P}$ = 5.4 Hz, 2), 19.28 (d, $J_{2\text{-}P}$ = 5.9 Hz, 2), 16.08 (d, $J_{\text{CH3-P}}$ = 7.4 Hz, OCH$_2$**CH$_3$**), 16.04 (d, $J_{\text{CH3-P}}$ = 6.9 Hz, OCH$_2$**CH$_3$**). ^{31}**P-NMR** (162 MHz, CDCl$_3$, 25 °C) δ 8.61 and 7.87 ppm. **HR-MS** (ESI) calculated for C$_{21}$H$_{27}$O$_9$N$_2$NaP 505.13464, found [M + Na]$^+$ 505.13456.

4,5-dimethoxy-2-nitrobenzyl (2S)-2-((ethoxy(pyrimidin-2-ylamino)phosphoryl)oxy) **propanoate** **(13)** was prepared from **24** (142.6 mg, 0.5 mmol, 1.0 eq.) and **20** (63.2 µL, 0.5 mmol, 1.0 eq.) according to the general procedure and purified by Flash silica gel chromatography using a gradient (CH$_2$Cl$_2$/MeOH 100:0→90:10, v/v), followed by reverse-phase chromatography using a gradient (H$_2$O/MeCN 100:0→0:100, v/v), which afforded **13** as a dense yellow oil (67.0 mg, 29%). ^{1}**H-NMR** (400 MHz, CDCl$_3$, 25 °C, 2 diastereoisomers in ca. 1:1 ratio) δ 8.53 (d, 1H, $J_{2'\text{-}P}$ = 4.9 Hz, 2′), 8.47 (d, 1H, $J_{2'\text{-}P}$ = 4.9 Hz, 2′), 7.73 and 7.71 (s, 2H, 3″), 7.13 and 7.07 (s, 2H, 6″), 6.88 (t, 2H, $J_{3'\text{-}P}$ = 4.9 Hz, 3′), 6.85 (t, 2H, $J_{3'\text{-}P}$ = 4.9 Hz, 3′), 5.64 (dd, 1H, J_{Gem} = 15.0 Hz, $J_{4a\text{-}6'}$ = 0.5 Hz, 4a), 5.59 (dd, 1H, J_{Gem} = 15.0 Hz, $J_{4b\text{-}6''}$ = 0.5 Hz, 4b), 5.56 (dd, 1H, J_{Gem} = 14.9 Hz, $J_{4a\text{-}6''}$ = 0.5 Hz, 4a), 5.51 (dd, 1H, J_{Gem} = 14.9 Hz, $J_{4b\text{-}6''}$ = 0.5 Hz, 4b), 5.43–5.29 (m, 2H, 1), 4.39–4.18 (m, 4H, OCH$_2$CH$_3$), 4.02 and 3.98 (s, 6H, 5″-O-**CH$_3$**), 3.96 and 3.96 (s, 6H, 4″-OCH$_3$), 1.66 (dd, 3H, $J_{2\text{-}1}$ = 6.9 Hz,

$J_{2\text{-}P}$ = 0.4 Hz, 2), 1.59 (dd, 3H, $J_{2\text{-}1}$ = 7.0 Hz, $J_{2\text{-}P}$ = 0.5 Hz, 2), 1.35 (td, 3H, $J_{CH3\text{-}CH2}$ = 7.1 Hz, $J_{CH3\text{-}P}$ = 0.9 Hz, OCH$_2$**CH$_3$**), 1.34 (td, 3H, $J_{CH3\text{-}CH2}$ = 7.1 Hz, $J_{CH3\text{-}P}$ = 0.9 Hz, OCH$_2$**CH$_3$**). **^{13}C-NMR** (101 MHz, CDCl$_3$, 25 °C) δ 170.73 (d, $J_{3\text{-}P}$ = 3.9 Hz, 3), 170.40 (d, $J_{3\text{-}P}$ = 4.6 Hz, 3), 159.22 (d, $J_{1'\text{-}P}$ = 6.2 Hz, 1'), 159.11 (d, $J_{1'\text{-}P}$ = 5.9 Hz, 1'), 158.38 (d, $J_{2'\text{-}P}$ = 1.5 Hz, 2'), 158.30 (d, $J_{2'\text{-}P}$ = 1.5 Hz, 2'), 153.81 and 153.65 (5"), 148.28 and 148.22 (4"), 139.67 and 139.56 (2"), 126.70 and 126.27 (1"), 114.56 and 114.51 (3'), 110.28 and 110.11 (6"), 108.12 and 108.11 (3"), 72.29 (d, $J_{1\text{-}P}$ = 5.4 Hz, 1), 72.19 (d, $J_{1\text{-}P}$ = 5.1 Hz, 1), 64.06 and 64.03 (4), 63.94 (d, $J_{CH2\text{-}P}$ = 5.5 Hz, O**CH$_2$**CH$_3$), 56.76 and 56.59 (5"-**OCH$_3$**), 56.39–56.32 (m, 4"-**OCH$_3$**), 19.37 (d, $J_{2\text{-}P}$ = 5.9 Hz, 2), 19.05 (d, $J_{2\text{-}P}$ = 6.7 Hz, 2), 16.07 (d, $J_{CH3\text{-}P}$ = 7.1 Hz, O**CH$_2$CH$_3$**), 16.03 (d, $J_{CH3\text{-}P}$ = 7.1 Hz, OCH$_2$**CH$_3$**). **^{31}P-NMR** (162 MHz, CDCl$_3$, 25 °C) δ −0.87 ppm. **HR-MS** (ESI) calculated for C$_{18}$H$_{24}$O$_9$N$_4$P 471.12754, found [M + H]$^+$ 471.12723.

4,5-dimethoxy-2-nitrobenzyl (2S)-2-((phenoxy(phenethylamino)phosphoryl)oxy)propanoate (**14**) was prepared from **24** (142.6 mg, 0.5 mmol, 1.0 eq.) and **21** (74.7 μL, 0.5 mmol, 1.0 eq.) according to the general procedure and purified by Flash silica gel chromatography using a gradient (hexane/EtOAc 100:0→0:100, v/v), followed by reverse-phase chromatography using a gradient (H$_2$O/MeCN 100:0→50:50, v/v), which afforded **14** as a dense yellow oil (79.4 mg, 29%). **^{1}H-NMR** (400 MHz, CDCl$_3$, 25 °C, 2 diastereoisomers in ca. 1:0.8 ratio) δ 7.74 and 7.71 (s, 2H, 3"), 7.38–7.11 (m, 20H, 2', 3', 4', 2''', 3''', 4'''), 7.04 and 7.02 (s, 2H, 6"), 5.63 (dd, J_{Gem} = 14.8 Hz, $J_{4a\text{-}6''}$ = 0.5 Hz, 1H, 4a), 5.63 (dd, 1H, J_{Gem} = 14.7 Hz, $J_{4a\text{-}6''}$ = 0.5 Hz, 4a), 5.58 (dd, 1H, J_{Gem} = 14.7 Hz, $J_{4b\text{-}6''}$ = 0.5 Hz, 4b), 5.55 (dd, 1H, J_{Gem} = 14.8 Hz, $J_{4b\text{-}6''}$ = 0.5 Hz, 4b), 5.11–5.01 (m, 2H, 1), 3.98 (s, 3H, 4"-**OCH$_3$**), 3.97 (s, 3H, 5"-**OCH$_3$**), 3.96 (s, 3H, 4"-**OCH$_3$**), 3.89 (s, 3H, 5"-**OCH$_3$**), 3.39–3.23 (m, 4H, NH**CH$_2$**), 2.96–2.84 (m, 2H, NH), 2.80–2.74 (m, 4H, **CH$_2$**Ph), 1.64 (d, 3H, $J_{2\text{-}1}$ = 6.9 Hz, 2), 1.57 (dd, 3H, $J_{2\text{-}1}$ = 7.0 Hz, $J_{2\text{-}P}$ = 0.3 Hz, 2). **^{13}C-NMR** (101 MHz, CDCl$_3$, 25 °C) δ 170.37 (d, $J_{3\text{-}P}$ = 4.5 Hz, 3), 170.15 (d, $J_{3\text{-}P}$ = 4.3 Hz, 3), 153.72 and 153.71 (5"), 150.8* (1'''), 148.44 and 148.32 (4"), 139.87 and 139.77 (2"), 138.29 and 138.18 (1'), 129.69 and 129.62 (3'), 128.86 and 128.84 (2'), 128.71 and 128.63 (3'''), 126.71 and 126.60 (4'''), 126.45 and 126.30 (1"), 124.97 and 124.92 (4'), 124.97 and 124.92 (2'), 120.27 (d, $J_{2'''\text{-}P}$ = 4.9 Hz, 2'''), 120.11 (d, $J_{2'''\text{-}P}$ = 5.1 Hz, 2'''), 110.40 and 110.35 (6"), 108.28 and 108.16 (3"), 71.27 (d, $J_{1\text{-}P}$ = 4.9 Hz, 1), 71.25 (d, $J_{1\text{-}P}$ = 5.3 Hz, 1), 64.22 and 64.12 (4), 56.64, 56.56, 56.44, and 56.41 (4"-**OCH$_3$**, 5"-**OCH$_3$**), 42.77 and 42.66 (NH**CH$_2$**), 37.64 (d, $J_{CH2\text{-}P}$ = 5.9 Hz, **CH$_2$**Ph), 37.61 (d, $J_{CH2\text{-}P}$ = 6.1 Hz, **CH$_2$**Ph), 19.44 (d, $J_{2\text{-}P}$ = 5.5 Hz, 2) 19.21 (d, $J_{2\text{-}P}$ = 5.8 Hz, 2). * The ^{13}C chemical shift was extracted from HMBC. **^{31}P-NMR** (162 MHz, CDCl$_3$, 25 °C) δ 4.40 and 3.55 ppm. **HR-MS** (ESI) calculated for C$_{26}$H$_{29}$O$_9$N$_2$NaP 567.15029, found [M + Na]$^+$ 567.14944.

4,5-dimethoxy-2-nitrobenzyl (2S)-2-((phenoxy(piperidin-1-yl)phosphoryl)oxy)propanoate (**15**) was prepared from **24** (142.6 mg, 0.5 mmol, 1.0 eq.) and **21** (74.7 μL, 0.5 mmol, 1.0 eq.) according to the general procedure and purified by Flash silica gel chromatography using a gradient (CH$_2$Cl$_2$/MeOH 100:0→90:10, v/v), followed by two reverse-phase chromatographies using a gradient (H$_2$O/MeCN 100:0→60:40, v/v), which afforded **15** as a dense yellow oil (75.5 mg, 30%). **^{1}H-NMR** (400 MHz, CDCl$_3$, 25 °C, 2 diastereoisomers in ca. 1:0.7 ratio) δ 7.75 and 7.72 (s, 2H, 3"), 7.34 and 7.27 (m, 4H, 3'''), 7.23 and 7.18 (m, 4H, 2'''), 7.17–7.10 (m, 2H, 4'''), 7.09 and 7.05 (s, 2H, 6"), 5.67–5.55 (m, 4H, 4), 5.12–5.01 (m, 2H, 1), 4.00 (s, 3H, 5"-**OCH$_3$**), 3.98 (s, 3H, 4"-**OCH$_3$**), 3.96 (s, 3H, 4"-**OCH$_3$**), 3.88 (s, 3H, 5"-**OCH$_3$**), 3.24–3.12 (m, 8H, 1'), 1.66 (d, 3H, $J_{2\text{-}1}$ = 6.9 Hz, 2), 1.58 (d, 3H, $J_{2\text{-}1}$ = 6.9 Hz, 2), 1.59–1.52 (m, 4H, 3'), 1.52–1.40 (m, 8H, 2'). **^{13}C-NMR** (101 MHz, CDCl$_3$, 25 °C) δ 170.32 (d, $J_{3\text{-}P}$ = 4.7 Hz, 3), 170.27 (d, $J_{3\text{-}P}$ = 4.6 Hz, 3), 153.77 and 153.75 (5"), 150.94 (d, $J_{1'''\text{-}P}$ = 6.9 Hz, 1'''), 150.86 (d, $J_{1'''\text{-}P}$ = 6.6 Hz, 1'''), 148.36 and 148.25 (4"), 139.78 and 139.68 (2"), 129.60 and 129.53 (3'''), 126.66 and 126.59 (1"), 124.71 (d, $J_{4'''\text{-}P}$ = 1.5 Hz, 4'''), 124.65 (d, $J_{4'''\text{-}P}$ = 1.5 Hz, 4'''), 120.17 (d, $J_{2'''\text{-}P}$ = 4.6 Hz, 2'''), 120.02 (d, $J_{2'''\text{-}P}$ = 5.4 Hz, 2'''), 110.28 and 110.26 (6"), 108.23 and 108.12 (3"), 71.06 (d, $J_{1'''\text{-}P}$ = 5.4 Hz, 1'''), 70.86 (d, $J_{1'''\text{-}P}$ = 5.3 Hz, 1'''), 64.07 and 64.01 (4), 56.68 and 56.58 (5"-**OCH$_3$**), 56.44 and 56.40 (4"-**OCH$_3$**), 45.51 (d, $J_{1'\text{-}P}$ = 2.3 Hz, 1'), 45.50 (d, $J_{1'\text{-}P}$ = 2.2 Hz, 1'), 25.79 (d, $J_{2'\text{-}P}$ = 4.6 Hz, 2'), 25.68 (d, $J_{2'\text{-}P}$ = 4.7 Hz, 2'), 24.24–24.20 (m, 3'), 19.48 (d, $J_{2\text{-}P}$ = 5.4 Hz, 2), 19.26 (d, $J_{2\text{-}P}$ = 5.4 Hz, 2). **^{31}P-NMR** (162 MHz,

CDCl$_3$, 25 °C) δ 3.96 and 3.47 ppm. **HR-MS** (ESI) calculated for C$_{23}$H$_{30}$O$_9$N$_2$P 509.16834, found [M + H]$^+$ 509.16852.

4,5-dimethoxy-2-nitrobenzyl (2S)-2-((phenoxy(phenylamino)phosphoryl)oxy)propanoate (**16**) was prepared from **24** (142.6 mg, 0.5 mmol, 1.0 eq.) and **21** (74.7 µL, 0.5 mmol, 1.0 eq.) according to the general procedure and purified by Flash silica gel chromatography using a gradient (hexane/EtOAc 100:0→0:100, *v/v*), followed by two reverse-phase chromatographies using a gradient (H$_2$O/MeCN 100:0→60:40, *v/v*), which afforded **16** as a dense yellow oil (53.5 mg, 21%). **^{1}H-NMR** (400 MHz, CDCl$_3$, 25 °C, 2 diastereoisomers in ca. 2:1 ratio) δ 7.73 and 7.71 (s, 2H, 3″), 7.32–6.96 (m, 20H, 2′, 3′, 4′, 2‴, 3‴, 4‴), 5.91 (d, 1H, $J_{NH\text{-}P}$ = 9.8 Hz, NH), 5.86 (d, 1H, $J_{NH\text{-}P}$ = 9.1 Hz, NH), 5.67 (dd, 1H, J_{Gem} = 14.8 Hz, $J_{4a\text{-}6″}$ = 0.6 Hz, 4a), 5.58 (dd, 1H, J_{Gem} = 14.8 Hz, $J_{4b\text{-}6″}$ = 0.6 Hz, 4b), 5.55 (dd, J_{Gem} = 14.5 Hz, $J_{4a\text{-}6″}$ = 0.6 Hz, 4a), 5.51 (dd, J_{Gem} = 14.5 Hz, $J_{4b\text{-}6″}$ = 0.6 Hz, 4b), 5.26–5.14 (m, 2H, 1), 3.97–3.96 (m, 6H, 4″-**OCH$_3$**), 3.90 and 3.87 (s, 6H, 5″-**OCH$_3$**), 1.67 (dd, 3H, $J_{2\text{-}1}$ = 6.9 Hz, $J_{2\text{-}P}$ = 0.5 Hz, 2), 1.60 (dd, 3H, $J_{2\text{-}1}$ = 6.9 Hz, $J_{2\text{-}P}$ = 0.3 Hz, 2). **^{13}C-NMR** (101 MHz, CDCl$_3$, 25 °C) δ 170.03 (d, $J_{3\text{-}P}$ = 5.0 Hz, 3), 169.75 (d, $J_{3\text{-}P}$ = 4.7 Hz, 3), 153.73 and 153.65 (5″), 150.29 (d, $J_{1‴\text{-}P}$ = 6.8 Hz, 1‴), 150.12 (d, $J_{1‴\text{-}P}$ = 6.5 Hz, 1‴), 148.42 and 148.34 (4″), 139.84 and 139.75 (2″), 138.63 and 138.58 (1′), 129.72 and 129.66 (3‴), 129.40 and 129.29 (3′), 126.37 and 126.06 (1″), 125.40–125.31 (4‴), 122.57 and 122.52 (4′), 120.43 (d, $J_{2‴\text{-}P}$ = 4.6 Hz, 2‴), 120.28 (d, $J_{2‴\text{-}P}$ = 4.7 Hz, 2‴), 118.24 (d, $J_{2′\text{-}P}$ = 7.3 Hz, 2′), 118.07 (d, $J_{2′\text{-}P}$ = 7.5 Hz, 2′), 110.39 and 110.33 (6″), 108.23 and 108.18 (3″), 72.00 (d, $J_{1\text{-}P}$ = 4.9 Hz, 1), 71.75 (d, $J_{1\text{-}P}$ = 4.8 Hz, 1), 64.37 and 64.30 (4), 56.54 and 56.51 (5″-**OCH$_3$**), 56.45–56.39 (4″-**OCH$_3$**), 19.23 (d, $J_{2\text{-}P}$ = 5.8 Hz, 2). **^{31}P-NMR** (162 MHz, CDCl$_3$, 25 °C) δ −2.50 and −2.99 ppm. **HR-MS** (ESI) calculated for C$_{24}$H$_{26}$O$_9$N$_2$P 517.13704, found [M + H]$^+$ 517.13695.

Methyl-1-cyclopropyl-7-(4-(((1-((4,5-dimethoxy-2-nitrobenzyl)oxy)-1-oxopropan-2-yl)oxy)(ethoxy)phosphoryl)piperazin-1-yl)-6-fluoro-4-oxo-1,4-dihydroquinoline-3-carboxylate (**17**). To a solution of **24** (71.3 mg, 0.25 mmol, 1 eq.) in 1.5 mL of dry toluene, **20** (31.6 µL, 0.25 mmol, 1.0 eq.), and TEA (45.4 µL, 0.325 mmol, 1.3 eq.) were added, and the reaction mixture was stirred at 25 °C for 16 h. The formation of the intermediate **25** was confirmed by ^{31}P NMR (δ$_P$ 4.6 and 4.1 ppm) before adding amine **27** (86.3 mg, 0.25 mmol, 1.0 eq.) and TEA (34.9 µL, 0.25 mmol, 1.0 eq.). The reaction mixture was stirred at room temperature until completion (1 h). After solvent evaporation, the crude residue was purified by Flash silica gel chromatography using a gradient (hexane/EtOAc/MeOH 50:50:0→0:100:0→0:50:50 *v/v/v*), which afforded **17** as a white solid (78.1 mg, 43%). **^{1}H-NMR** (400 MHz, CDCl$_3$, 25 °C, 2 diastereoisomers in ca. 1:1 ratio) δ 8.53 and 8.53 (s, 2H, 9′), 8.00 (d, 1H, $J_{7′\text{-}F}$ = 13.2 Hz, 7′), 8.00 (d, 1H, $J_{7′\text{-}F}$ = 13.1 Hz, 7′), 7.72 and 7.67 (s, 2H, 3″), 7.27 (d, 1H, $J_{4′\text{-}F}$ = 7.2 Hz, 4′), 7.26 (d, 1H, $J_{4′\text{-}F}$ = 7.1 Hz, 4′), 7.05 and 7.02 (s, 2H, 6″), 5.66–5.53 (m, 4H, 4), 5.06–4.92 (m, 2H, 1), 4.19–4.06 (m, 4H, **OCH$_2$CH$_3$**), 4.01 and 4.00 (s, 6H, 5″-**OCH$_3$**), 3.97 and 3.93 (s, 6H, 4″-**OCH$_3$**), 3.91–3.90 (m, 6H, 16′), 3.50–3.30 (m, 10H, 1′, 12′), 3.26–3.08 (m, 8H, 2′), 1.65–1.60 (m, 6H, 2), 1.43–1.28 (m, 10H, OCH$_2$**CH$_3$**, 13′ or 14′), 1.18–1.11 (m, 4H, 13′ or 14′). **^{13}C-NMR** (101 MHz, CDCl$_3$, 25 °C) δ 173.04 (d, $J_{11′\text{-}F}$ = 2.1 Hz, 11′), 170.00 (d, $J_{11′\text{-}F}$ = 2.2 Hz, 11′), 170.50 (d, $J_{3\text{-}P}$ = 4.6 Hz, 3), 170.43 (d, $J_{3\text{-}P}$ = 5.4 Hz, 3), 166.40 and 166.35 (15′), 153.38 (d, $J_{8′\text{-}F}$ = 248.6 Hz, 8′), 153.63 (5″), 148.58 and 148.55 (4″), 148.42 (9′), 144.55 (d, $J_{3′\text{-}F}$ = 10.5 Hz, 3′), 144.49 (d, $J_{3′\text{-}F}$ = 10.6 Hz, 3′), 140.06 (2″), 138.01 (d, $J_{5′\text{-}F}$ = 1.5 Hz, 5′), 137.99 (d, $J_{5′\text{-}F}$ = 1.5 Hz, 5′), 126.09 and 125.97 (1″), 123.36 (d, $J_{6′\text{-}F}$ = 7.3 Hz, 6′), 123.28 (d, $J_{6′\text{-}F}$ = 6.9 Hz, 6′), 113.40 (d, $J_{7′\text{-}F}$ = 23.5 Hz, 7′), 113.30 (d, $J_{7′\text{-}F}$ = 22.9 Hz, 7′), 110.94 and 110.85 (6″), 110.12 and 110.07 (10′), 108.30 and 108.26 (3″), 105.13 (d, $J_{4′\text{-}F}$ = 9.5 Hz, 4′), 105.12 (d, $J_{4′\text{-}F}$ = 9.1 Hz, 4′), 70.69 (d, $J_{1\text{-}F}$ = 4.8 Hz, 1), 70.65 (d, $J_{1\text{-}F}$ = 5.6 Hz, 1), 64.21 and 64.18 (4), 63.21 (d, $J_{CH2\text{-}P}$ = 6.0 Hz, OCH$_2$CH$_3$), 63.00 (d, $J_{CH2\text{-}P}$ = 6.0 Hz, OCH$_2$CH$_3$), 56.68 and 56.63 (5″-**OCH$_3$**), 56.45 and 56.44 (4″-**OCH$_3$**), 52.07 and 52.06 (16′), 50.26 and 50.09 (m, 2′), 44.40 and 44.38 (1′), 34.56 and 34.55 (12′), 19.50 (d, $J_{2\text{-}P}$ = 5.2 Hz, 2), 16.18 (d, $J_{CH3\text{-}P}$ = 7.0 Hz, OCH$_2$**CH$_3$**), 16.14 (d, $J_{CH3\text{-}P}$ = 7.0 Hz, OCH$_2$**CH$_3$**), 8.15 (m, 13′, 14′). **^{31}P-NMR** (162 MHz, CDCl$_3$, 25 °C) δ 7.87 and 7.17 ppm. **HR-MS** (ESI) calculated for C$_{32}$H$_{39}$O$_{12}$N$_4$FP 721.22806, found [M + H]$^+$ 721.22796.

Methyl-1-cyclopropyl-7-(4-(ethoxy((1-ethoxy-1-oxopropan-2-yl)oxy)phosphoryl) piperazin-1-yl)-6-fluoro-4-oxo-1,4-dihydroquinoline-3-carboxylate (**18**). To a solution of (-)-Ethyl L-Lactate (28.7 µL, 0.25 mmol, 1 eq.) in 1.5 mL of dry toluene, **20** (31.6 µL, 0.25 mmol, 1.0 eq.) and TEA (45.4 µL, 0.325 mmol, 1.3 eq.) were added, and the reaction mixture was stirred at 25 °C for 16 h. The formation of the intermediate was confirmed by ^{31}P-NMR (δ_P 4.7 and 3.7 ppm) before adding amine **27** (86.3 mg, 0.25 mmol, 1.0 eq.) and TEA (34.9 µL, 0.25 mmol, 1.0 eq.). The reaction mixture was stirred at 25 °C until completion (1 h). After evaporating the solvent, the crude residue was purified by Flash silica gel chromatography using a gradient (hexane/EtOAc 50:50→0:100, v/v), followed by reverse-phase chromatography using a gradient (H$_2$O/MeCN 95:5→50:50, v/v), which afforded **18** as white solid (61.1 mg, 44%). ^{1}H-NMR (400 MHz, CDCl$_3$, 25 °C, 2 diastereoisomers in ca. 1:1 ratio) δ 8.60–8.57 (m, 2H, 9'), 8.08 (d, 1H, $J_{7'-F}$ = 13.2 Hz, 7'), 8.07 (d, 1H, $J_{7'-F}$ = 13.2 Hz, 7'), 7.31–7.27 (m, 2H, 4'), 4.95–4.83 (m, 2H, 1), 4.30–4.08 (m, 8H, 4, OCH$_2$CH$_3$), 3.95–3.93 (m, 6H, 16'), 3.53–3.36 (m, 10H, 1', 12'), 3.28–3.20 (m, 8H, 2'), 1.60 (d, 3H, J_{2-P} = 7.0 Hz, 2), 1.57 (d, 3H, J_{2-P} = 6.9 Hz, 2), 1.40–1.29 (m, 16H, 13' or 14', 5, OCH$_2$CH$_3$), 1.19–1.13 (m, 4H, 13' or 14'). ^{13}C-NMR (101 MHz, CDCl$_3$, 25 °C) δ 173.17–173.02 (m, 11'), 171.00 (d, J_{3-P} = 4.5 Hz, 3), 170.91 (d, J_{3-P} = 5.4 Hz, 3), 166.57–166.49 (m, 15'), 153.44 (d, $J_{8'-F}$ = 249.1 Hz, 8'), 148.45 (9'), 144.76–144.55 (m, 3'), 138.02 (5'), 123.63–123.24 (m, 6'), 113.50 (d, $J_{7'-F}$ = 23.1 Hz, 7'), 113.47 (d, $J_{7'-F}$ = 22.8 Hz, 7'), 110.32–110.05 (m, 10'), 105.05–104.96 (m, 4'), 70.78 (d, J_{1-P} = 5.4 Hz, 1), 63.06 (d, J_{CH2-P} = 6.1 Hz, OCH$_2$CH$_3$), 62.94 (d, J_{CH2-P} = 6.2 Hz, OCH$_2$CH$_3$), 61.46 and 61.41 (4), 52.14 (16'), 50.41–50.14 (m, 2'), 44.43–44.34 (m, 1'), 34.54 (12'), 19.54–19.36 (2), 16.20 (d, J_{CH3-P} = 6.9 Hz, OCH$_2$CH$_3$), 16.17 (d, J_{CH3-P} = 6.6 Hz, OCH$_2$CH$_3$), 14.21 and 14.15 (5), 8.18 (m, 13', 14'). ^{31}P-NMR (162 MHz, CDCl$_3$, 25 °C) δ 7.78 and 6.98 ppm. HR-MS (ESI) calculated for C$_{25}$H$_{34}$O$_8$N$_3$FP 554.20621, found [M + H]$^+$ 554.20614.

Methyl-1-cyclopropyl-6-fluoro-4-oxo-7-(piperazin-1-yl)-1,4-dihydroquinoline-3-carboxylate (**27**). To a suspension of Ciprofloxacin (1.0 g, 3.0 mmol, 1 eq.) in 30 mL of dry methanol, cooled to 0 °C, SOCl$_2$ was added dropwise, and the resulting solution was refluxed for 16 h [33]. After evaporating the solvent, the crude residue was dissolved in CH$_2$Cl$_2$ and washed with a saturated solution of K$_2$CO$_3$, and the organic layer was dried over Na$_2$SO$_4$. After filtration and solvent removal, the crude residue was purified by Flash silica gel chromatography using a gradient (CH$_2$Cl$_2$/MeOH 100:0→70:30, v/v), to obtain methyl ester **27** as a white solid (0.78 g, 76%).

Characterization of intermediates (I), products (P) obtained after irradiation with UV light, and undesired products (X)

1-I. ^{13}C-NMR (126 MHz, 50% CACO/DMSO, 25 °C) δ 140.73 (1'), 130.43 (2' or 3'), 130.07 (3' or 2'), 127.93 (4'), 68.73 (d, J_{2-P} = 5.8 Hz, 2), 64.25 (d, J_{1-P} = 5.4 Hz, 1), 61.94 (d, J_{CH2-P} = 7.9 Hz, OCH$_2$CH$_3$), 43.81 (NH-CH$_2$), 38.71 (d, J_{CH2-P} = 6.1 Hz, CH$_2$Ph), 17.25 (d, J_{CH3-P} = 6.5 Hz, OCH$_2$CH$_3$). ^{31}P-NMR (202 MHz, 50% CACO/DMSO-d_6, 25 °C) δ 10.83 ppm. HR-MS (ESI-) calculated for C$_{12}$H$_{19}$O$_4$NP 272.10572, found [M−H]$^-$ 272.10551.

2-I. ^{13}C-NMR (126 MHz, 50% CACO/DMSO-d_6, 25 °C) δ 68.79 (d, J_{2-P} = 5.5 Hz, 2), 64.34 (d, J_{1-P} = 5.6 Hz, 1), 61.92 (d, J_{CH2-P} = 7.6 Hz, OCH$_2$CH$_3$), 46.33 (d, $J_{1'-P}$ = 2.1 Hz, 1'), 26.97 (d, $J_{2'-P}$ = 4.7 Hz, 2'), 25.15 (3'), 17.25 (d, J_{CH3-P} = 6.6 Hz, OCH$_2$CH$_3$). ^{31}P-NMR (202 MHz, 50% CACO/DMSO-d_6, 25 °C) δ 9.81 ppm. HR-MS (ESI−) calculated for C$_9$H$_{19}$O$_4$NP 236.10572, found [M−H]$^-$ 236.10585.

3-I. ^{13}C-NMR (126 MHz, 50% CACO/DMSO-d_6, 25 °C) δ 141.11 (1'), 131.01 (3'), 123.48 (4'), 119.19 (d, $J_{2'-P}$ = 7.6 Hz, 2'), 69.46 (d, J_{2-P} = 5.9 Hz, 2), 65.11 (d, J_{1-P} = 5.6 Hz, 1), 61.82 (d, J_{CH2-P} = 7.9 Hz, OCH$_2$CH$_3$), 17.19 (d, J_{CH3-P} = 6.9 Hz, OCH$_2$CH$_3$). ^{31}P-NMR (202 MHz, 50% CACO/DMSO-d_6, 25 °C) δ 3.62 ppm. HR-MS (ESI−) calculated for C$_{10}$H$_{15}$O$_4$NP 244.07442, found [M−H]$^-$ 244.07419.

4-I. ^{13}C-NMR (126 MHz, 50% HEPES/DMSO-d_6, 25 °C) δ 151.73 (d, $J_{1'''-P}$ = 6.8 Hz, 1'''), 140.51 (1'), 131.54 (3'''), 130.40 (2'), 130.08 (3'), 127.98 (4'), 126.77 (4'''), 121.67 (d, $J_{2'''-P}$ = 4.6 Hz, 2'''), 69.41 (d, J_{2-P} = 5.9 Hz, 2), 61.85 (d, J_{1-P} = 7.8 Hz, 1), 44.04 (NHCH$_2$), 38.58 (d, J_{CH2-P} = 6.1 Hz, CH$_2$Ph). ^{31}P-NMR (202 MHz, 50% HEPES/DMSO-d_6, 25 °C) δ 6.32 ppm. HR-MS (ESI) calculated for C$_{16}$H$_{20}$O$_4$NNaP 344.10222, found [M + Na]$^+$ 344.10212.

4-cyc-I. ^{13}**C-NMR** (126 MHz, 50% CACO/DMSO-d_6, 25 °C) δ 140.45 (1′), 130.41 (2′), 130.08 (3′), 127.98 (4′), 67.74—67.69 (m, 2, 1), 43.71 (NH**CH$_2$**), 38.67 (d, J_{CH2-P} = 5.2 Hz, **CH$_2$**Ph). ^{31}**P-NMR** (202 MHz, 50% CACO/DMSO-d_6, 25 °C) δ 28.06 ppm. **HR-MS** (ESI−) calculated for C$_{10}$H$_{13}$O$_3$NP 226.06385, found [M−H]$^-$ 226.06358.

4-P2. ^{13}**C-NMR** (126 MHz, 50% CACO/DMSO-d_6, 25 °C) δ 141.56 (1′), 130.38 (2′), 130.03 (3′), 127.69 (4′), 66.76 (d, J_{2-P} = 5.3 Hz, 2), 62.90 (d, J_{1-P} = 7.3 Hz, 1), 44.60 (NH**CH$_2$**), 39.2* (**CH$_2$**Ph). * The ^{13}C chemical shift was extracted from HMBC. ^{31}**P-NMR** (202 MHz, 50% CACO/DMSO-d_6, 25 °C) δ 8.05 ppm. **HR-MS** (ESI−) calculated for C$_{10}$H$_{15}$O$_4$NP 244.07442, found [M−H]$^-$ 244.07457.

5-I. ^{13}**C-NMR** (126 MHz, 50% CACO/DMSO-d_6, 25 °C) δ 151.72 (d, $J_{1'''-P}$ = 6.8 Hz, 1′′′), 131.58 (3′′′), 126.83 (4′′′), 121.64 (d, $J_{2'''-P}$ = 5.1 Hz, 2′′′), 64.48 (d, J_{2-P} = 6.0 Hz, 2), 61.84 (d, J_{1-P} = 7.9 Hz, 1), 46.53 (d, $J_{1'-P}$ = 1.9 Hz, 1′), 26.77 (d, $J_{2'-P}$ = 4.5 Hz, 2′), 25.01 (3′). ^{31}**P-NMR** (202 MHz, 50% CACO/DMSO-d_6, 25 °C) δ 5.10 ppm. **HR-MS** (ESI) calculated for C$_{13}$H$_{21}$O$_4$NP 286.12027, found [M + H]$^+$ 286.12039.

5-cyc-I. ^{13}**C-NMR** (126 MHz, 50% CACO/DMSO-d_6, 25 °C) δ 67.87—67.83 (m, 2, 1), 46.54 (d, $J_{1'-P}$ = 3.1 Hz, 1′), 27.03 (d, $J_{2'-P}$ = 3.6 Hz, 2′), 25.07 (3′). ^{31}**P-NMR** (202 MHz, 50% CACO/DMSO-d_6, 25 °C) δ 26.48 ppm. **HR-MS** (ESI) calculated for C$_7$H$_{15}$O$_3$NP 192.07841, found [M + H]$^+$ 192.07829.

6-I. ^{13}**C-NMR** (126 MHz, 50% CACO/DMSO-d_6, 25 °C) δ 151.39 (d, $J_{1'''-P}$ = 6.7 Hz, 1′′′), 140.63 (1′), 131.66 (3′′′), 131.10 (3′), 127.18 (4′′′), 123.92 (4′), 121.55 (d, $J_{2'''-P}$ = 4.7 Hz, 2′′′), 119.50 (d, $J_{2'-P}$ = 7.7 Hz, 2′), 70.19 (d, J_{2-P} = 5.8 Hz, 2), 61.76 (d, J_{1-P} = 7.6 Hz, 1). ^{31}**P-NMR** (202 MHz, 50% CACO/DMSO-d_6, 25 °C) δ −0.94 ppm.

6-cyc-I. ^{13}**C-NMR** (126 MHz, 50% CACO/DMSO-d_6, 25 °C) δ 140.23 (1′), 131.15 (3′), 124.73 (4′), 120.98 (d, $J_{2'-P}$ = 7.3 Hz, 2′), 68.02—67.95 (m, 2, 1). ^{31}**P-NMR** (202 MHz, 50% CACO/DMSO-d_6, 25 °C) δ 21.66 ppm.

9-I. HR-MS (ESI−) calculated for C$_{11}$H$_{15}$O$_5$NP 272.06933, found [M−H]$^-$ 272.06926.

13-I. ^{13}**C-NMR** (126 MHz, 50% CACO/DMSO-d_6, 25 °C) δ 177.54 (d, J_{3-P} = 4.5 Hz, 3), 160.21–160.10 (m, 2′), 160.05–159.93 (m, 1′), 116.69 and 116.67 (3′), 76.31 (d, J_{1-P} = 6.3 Hz, 1), 76.20 (d, J_{1-P} = 6.1 Hz, 1), 65.89 (d, J_{CH2-P} = 6.1 Hz, O**CH$_2$**CH$_3$), 65.87 (d, J_{CH2-P} = 6.2 Hz, O**CH$_2$**CH$_3$), 21.29 (d, J_{2-P} = 5.7 Hz, 2), 21.08 (d, J_{2-P} = 5.5 Hz, 2), 17.14 (d, J_{CH3-P} = 7.3 Hz, OCH$_2$**CH$_3$**), 17.10 (d, J_{CH3-P} = 6.7 Hz, OCH$_2$**CH$_3$**). ^{31}**P-NMR** (202 MHz, 50% CACO/DMSO-d_6, 25 °C) δ −0.13 and −0.75 ppm. **HR-MS** (ESI−) calculated for C$_9$H$_{13}$O$_5$N$_3$P 274.05983, found [M − H]$^-$ 274.05995.

16-I. ^{13}**C-NMR** (126 MHz, 50% CACO/DMSO-d_6, 25 °C) δ 153.93 (d, J_{3-P} = 6.8 Hz, 3), 151.68–151.46 (m, 1′′′), 140.89 (1′). 140.79 (1′), 131.55 (3′′′ or 3′), 131.53 (3′′′ or 3′), 131.03 (3′ or 3′′′), 130.98 (3′ or 3′′′), 127.03 (4′′′), 123.66 (4′), 123.63 (4′), 121.69–121.60 (m 2′′′), 119.36 (d, $J_{2'-P}$ = 8.0 Hz, 2′), 76.35 (d, J_{1-P} = 5.9 Hz, 1), 21.31 (d, J_{2-P} = 4.8 Hz, 2). ^{31}**P-NMR** (202 MHz, 50% CACO/DMSO-d_6, 25 °C) δ −2.21 and −2.63 ppm. **HR-MS** (ESI−) calculated for C$_{15}$H$_{15}$O$_5$NP 320.06933, found [M − H]$^-$ 320.06956.

7-P (= **8-P, 9-P, 10-P, 11-P, 12-P, 13-P, 17-P, 10-P, 11-P, 12-P, 13-P, 17-P, 18-P**). ^{13}**C-NMR** (126 MHz, 50% CACO/DMSO-d_6, 25 °C) δ 179.92 (d, J_{3-P} = 6.5 Hz, 3), 73.60 (d, J_{1-P} = 5.6 Hz, 1), 62.56 (d, J_{CH2-P} = 5.5 Hz, O**CH$_2$**CH$_3$), 21.60 (d, J_{2-P} = 3.1 Hz, 2), 17.52 (d, J_{CH3-P} = 7.4 Hz, OCH$_2$**CH$_3$**). ^{31}**P-NMR** (202 MHz, 50% CACO/DMSO-d_6, 25 °C) δ −1.09 ppm. **HR-MS** (ESI−) calculated for C$_5$H$_{10}$O$_6$P 197.02205, found [M − H]$^-$ 197.02187.

14-P1. ^{13}**C-NMR** (126 MHz, 50% CACO/DMSO-d_6, 25 °C) δ 179.52 (d, J_{3-P} = 7.1 Hz, 3), 153.92 (d, $J_{1'''-P}$ = 6.9 Hz, 1′′′), 130.95 (3′′′), 124.87 (4′′′), 121.76 (d, $J_{2'''-P}$ = 4.6 Hz, 2′′′), 74.33 (d, J_{1-P} = 6.1 Hz, 1), 21.51 (d, J_{2-P} = 2.9 Hz, 2). ^{31}**P-NMR** (202 MHz, 50% CACO/DMSO-d_6, 25 °C) δ −6.24 ppm. **HR-MS** (ESI−) calculated for C$_9$H$_{10}$O$_6$NP 245.02205, found [M−H]$^-$ 245.02184.

7-X. ^{13}**C-NMR** (126 MHz, 50% CACO/DMSO-d_6, 25 °C) δ 175.44 (d, J_{3-P} = 6.5 Hz, 3), 140.41 (1′), 130.32 (2′), 130.07 (3′), 127.96 (4′), 72.75 (d, J_{1-P} = 5.5 Hz, 1), 62.90–62.78 (m, O**CH$_2$**CH$_3$), 41.61 (NH-**CH$_2$**), 36.02 (**CH$_2$**Ph), 21.10–21.01 (m, 2), 17.61–17.45 (m, OCH$_2$**CH$_3$**). ^{31}**P-NMR** (202 MHz, 50% CACO/DMSO-d_6, 25 °C) δ −1.76 ppm. **HR-MS** (ESI−) calculated for C$_{13}$H$_{19}$O$_5$NP 300.10063, found [M − H]$^-$ 300.10069.

9-X. ^{13}**C-NMR** (126 MHz, 50% CACO/DMSO-d_6, 25 °C) δ 173.93 (d, $J_{3\text{-}P}$ = 6.2 Hz, 3), 138.54 (1′), 130.60 (3′), 126.60 (4′), 122.38 (2′), 73.05 (d, $J_{1\text{-}P}$ = 5.5 Hz, 1), 62.94 (d, $J_{CH2\text{-}P}$ = 5.8 Hz, O**CH$_2$**CH$_3$), 20.93 (d, $J_{2\text{-}P}$ = 3.3 Hz, 2), 17.52 (d, $J_{CH3\text{-}P}$ = 7.2 Hz, OCH$_2$**CH$_3$**). ^{31}**P-NMR** (202 MHz, 50% CACO/DMSO-d_6, 25 °C) δ −1.57 ppm. **HR-MS** (ESI−) calculated for C$_{11}$H$_{15}$O$_5$NP 272.06933, found [M−H]$^-$ 272.06943.

5. Conclusions

In summary, we designed and synthesized phosphate-based SI linkers for amine-containing drug delivery. We found that the lactate spacer can release amines effectively within 15 min; moreover, it can release two cargos sequentially—the first amine cargo within minutes and the second phenolic cargo overnight. Surprisingly, this is exactly the opposite release order that we found when using an ethylene glycol SI spacer, whereby phenol is released preferentially [22]. Interestingly, the linkers bearing primary amines lack stability at physiological pH (pH = 7.4) due to an intramolecular rearrangement caused by the nucleophilic attack of NH nitrogen from LG on the carbonyl group of lactate. This alternative decomposition, which yields the undesired product **X**, can be suppressed by changing the buffer (e.g., HEPES instead of Cacodylate buffer, pH = 7.4), by decreasing the buffer pH to mildly acidic (pH = 5), or by *N*-methylation of phosphoramidate nitrogen. In turn, derivatives bearing secondary amines are stable in a range of pH 5–7.4. As such, our prodrug approach is the most suitable for the delivery of secondary amines. Further applicability was demonstrated by phosphorylation of the antibiotic Ciprofloxacin, whose phototriggerable and enzyme-triggerable prodrugs released Ciprofloxacin successfully. Overall, our results establish an experimental paradigm for the smart design of new self-immolative systems for the targeted delivery of various amine-containing drugs and their enhanced cellular uptake and activity, thus broadening the applications of prodrug technology. Moreover, phospholane amidates could lead to the design of new synthetic approaches in phosphorus chemistry.

Sample Availability: Samples of the compounds are not available from the authors.

References

1. Barbosa-Filho, J.M.; Piuvezam, M.R.; Moura, M.D.; Silva, M.S.; Lima, K.V.B.; da-Cunha, E.V.L.; Fechine, I.M.; Takemura, O.S. Anti-inflammatory activity of alkaloids: A twenty-century review. *Rev. Bras. Farmacogn.* **2006**, *16*, 109–139. [CrossRef]
2. Ballout, F.; Habli, Z.; Monzer, A.; Rahal, O.N.; Fatfat, M.; Gali-Muhtasib, H. Anticancer Alkaloids: Molecular Mechanisms and Clinical Manifestations. In *Bioactive Natural Products for the Management of Cancer: From Bench to Bedside*; Springer: Singapore, 2019; pp. 1–35.
3. Mohan, K.; Jeyachandran, R. Alkaloids as anticancer agents. *Ann. Phytomed.* **2012**, *1*, 46–53.
4. Krishnan, N.; Devadasan, V.; Raman, P. Plant-derived alkaloids as anti-viral agents. *Int. J. Res. Pharm. Sci.* **2020**, *11*, 6174–6182. [CrossRef]
5. Cushnie, T.P.T.; Cushnie, B.; Lamb, A.J. Alkaloids: An overview of their antibacterial, antibiotic-enhancing and antivirulence activities. *Int. J. Antimicrob. Agents* **2014**, *44*, 377–386. [CrossRef]
6. Sayhan, H.; Beyaz, S.G.; Çeliktaş, A. The Local Anesthetic and Pain Relief Activity of Alkaloids. In *Alkaloids—Alternatives in Synthesis, Modification and Application*; InTechOpen: London, UK, 2017. [CrossRef]
7. Krise, J.P.; Oliyai, R. Prodrugs of Amines. In *Prodrugs*; Stella, V., Borchardt, R., Hageman, M., Oliyai, R., Maag, H., Tilley, J.E., Eds.; Springer: New York, NY, USA, 2008; pp. 801–831.
8. Albert, A. Chemical aspects of selective toxicity. *Nature* **1958**, *182*, 421–423. [CrossRef] [PubMed]
9. Abet, V.; Filace, F.; Recio, J.; Alvarez-Builla, J.; Burgos, C. Prodrug approach: An overview of recent cases. *Eur. J. Med. Chem.* **2017**, *127*, 810–827. [CrossRef] [PubMed]
10. Rautio, J.; Kumpulainen, H.; Heimbach, T.; Oliyai, R.; Oh, D.; Järvinen, T.; Savolainen, J. Prodrugs: Design and clinical applications. *Nat. Rev. Drug Discov.* **2008**, *7*, 255–270. [CrossRef] [PubMed]
11. Simplício, A.L.; Clancy, J.M.; Gilmer, J.F. Prodrugs for amines. *Molecules* **2008**, *13*, 519–547. [CrossRef]
12. Alouane, A.; Labruère, R.; Le Saux, T.; Schmidt, F.; Jullien, L. Self-immolative spacers: Kinetic aspects, structure-property relationships, and applications. *Angew. Chem.-Int. Ed.* **2015**, *54*, 7492–7509. [CrossRef]
13. Gonzaga, R.V.; do Nascimento, L.A.; Santos, S.S.; Machado Sanches, B.A.; Giarolla, J.; Ferreira, E.I. Perspectives About Self-Immolative Drug Delivery Systems. *J. Pharm. Sci.* **2020**, *109*, 3262–3281. [CrossRef] [PubMed]
14. Teicher, B.A.; Chari, R.V.J. Antibody conjugate therapeutics: Challenges and potential. *Clin. Cancer Res.* **2011**, *17*, 6389–6397. [CrossRef] [PubMed]
15. Egron, D.; Lefebvre, I.; Périgaud, C.; Beltran, T.; Pompon, A.; Gosselin, G.; Aubertin, A.M.; Imbach, J.L. Anti-HIV pronucleotides: Decomposition pathways and correlation with biological activities. *Bioorg. Med. Chem. Lett.* **1998**, *8*, 1045–1050. [CrossRef]
16. Wei, Y.; Qiu, G.; Lei, B.; Qin, L.; Chu, H.; Lu, Y.; Zhu, G.; Gao, Q.; Huang, Q.; Qian, G.; et al. Oral Delivery of Propofol with Methoxymethylphosphonic Acid as the Delivery Vehicle. *J. Med. Chem.* **2017**, *60*, 8580–8590. [CrossRef] [PubMed]
17. Mehellou, Y.; Rattan, H.S.; Balzarini, J. The ProTide Prodrug Technology: From the Concept to the Clinic. *J. Med. Chem.* **2018**, *61*, 2211–2226. [CrossRef]
18. Ray, A.S.; Fordyce, M.W.; Hitchcock, M.J.M. Tenofovir alafenamide: A novel prodrug of tenofovir for the treatment of Human Immunodeficiency Virus. *Antivir. Res.* **2016**, *125*, 63–70. [CrossRef] [PubMed]
19. McQuaid, T.; Savini, C.; Seyedkazemi, S. Sofosbuvir, a significant paradigm change in hcv treatment. *J. Clin. Transl. Hepatol.* **2015**, *3*, 27–35.
20. Siegel, D.; Hui, H.C.; Doerffler, E.; Clarke, M.O.; Chun, K.; Zhang, L.; Neville, S.; Carra, E.; Lew, W.; Ross, B.; et al. Discovery and Synthesis of a Phosphoramidate Prodrug of a Pyrrolo[2,1-f][triazin-4-amino]Adenine C-Nucleoside (GS-5734) for the Treatment of Ebola and Emerging Viruses. *J. Med. Chem.* **2017**, *60*, 1648–1661. [CrossRef]
21. Klán, P.; Šolomek, T.; Bochet, C.G.; Blanc, A.; Givens, R.; Rubina, M.; Popik, V.; Kostikov, A.; Wirz, J. Photoremovable protecting groups in chemistry and biology: Reaction mechanisms and efficacy. *Chem. Rev.* **2013**, *113*, 119–191. [CrossRef]
22. Procházková, E.; Šimon, P.; Straka, M.; Filo, J.; Majek, M.; Cigáň, M.; Baszczyński, O. Phosphate Linkers with Traceable Cyclic Intermediates for Self-Immolation Detection and Monitoring. *Chem. Commun.* **2020**, *57*, 211–214. [CrossRef]
23. Zhang, S.; Wang, H.; Shen, Y.; Zhang, F.; Seetho, K.; Zou, J.; Taylor, J.S.A.; Dove, A.P.; Wooley, K.L. A simple and efficient synthesis of an acid-labile polyphosphoramidate by organobase-catalyzed ring-opening polymerization and transformation to polyphosphoester ionomers by acid treatment. *Macromolecules* **2013**, *46*, 5141–5149. [CrossRef]
24. Procházková, E.; Filo, J.; Cigáň, M.; Baszczyński, O. Sterically-Controlled Self-Immolation in Phosphoramidate Linkers Triggered by Light. *Eur. J. Org. Chem.* **2020**, *2020*, 897–906. [CrossRef]
25. Šimon, P.; Tichotová, M.; Gallardo, M.G.; Procházková, E.; Baszczynski, O. Phosphate-Based Self-Immolative Linkers for Tunable Double Cargo Release. *Chem. A Eur. J.* in press. [CrossRef]
26. Mulliez, M.; Wolf, R. Contraste entre la phosphorylation des alcools et des amines par les 2-4 dioxo, oxa-1, aza-3 phospholanes-2. *Bull. Soc. Chim. Fr.* **1986**, 101–108.
27. Jafar, N.N.A.; Majeed, N.S. Microwave-assisted synthesis and biological activity of ester, carbothioate and carbohydrazide derivative compounds of the drug Ciprofloxacin. *J. Chem. Pharm. Sci.* **2017**, *10*, 515–521.
28. Albert, A.; Goldacre, R.; Phillips, J. 455. The strength of heterocyclic bases. *J. Chem. Soc.* **1948**, 2240–2249. [CrossRef]
29. Dudkin, S.M.; Ledneva, R.K.; Shabarova, Z.A.; Prokofiev, M.A. Hydrolysis of uridine-5′ N-aryl and N-alkyl phosphoramidates by ribonucleoside-5′ phosphoramidase. *FEBS Lett.* **1971**, *16*, 48–50. [CrossRef]
30. Beltran, T.; Egron, D.; Pompon, A.; Lefebvre, I.; Périgaud, C.; Gosselin, G.; Aubertin, A.M.; Imbach, J.L. Rational design of a new series of pronucleotide. *Bioorg. Med. Chem. Lett.* **2001**, *11*, 1775–1777. [CrossRef]

31. Drontle, D. Designing a Pronucleotide Stratagem: Lessons from Amino Acid Phosphoramidates of Anticancer and Antiviral Pyrimidines. *Mini Rev. Med. Chem.* **2004**, *4*, 409–419. [CrossRef]
32. Kanzian, T.; Nigst, T.A.; Maier, A.; Pichl, S.; Mayr, H. Nucleophilic Reactivities of Primary and Secondary Amines in Acetonitrile. *Eur. J. Org. Chem.* **2009**, 6379–6638. [CrossRef]
33. Sachin, K.; Kim, E.M.; Cheong, S.J.; Jeong, H.J.; Lim, S.T.; Sohn, M.H.; Kim, D.W. Synthesis of N 4'-[18F]fluoroalkylated ciprofloxacin as a potential bacterial infection imaging agent for PET study. *Bioconjug. Chem.* **2010**, *21*, 2282–2288. [CrossRef]

III.

M. Tichotová, A. Ešnerová, L. Tučková, L. Bednárová, I. Císařová, O. Baszczyňski and E. Procházková:

^{31}P NMR Parameters May Facilitate the Stereochemical Analysis of Phosphorus-Containing Compounds

JMR

^{31}P NMR parameters may facilitate the stereochemical analysis of phosphorus-containing compounds

Markéta Tichotová [a,b], Aneta Ešnerová [c], Lucie Tučková [a], Lucie Bednárová [a], Ivana Císařová [d], Ondřej Baszczyński [a,c], Eliška Procházková [a,*]

[a] Institute of Organic Chemistry and Biochemistry, Czech Academy of Sciences, 160 00 Prague, Czech Republic
[b] Department of Physical and Macromolecular Chemistry, Faculty of Science, Charles University, 116 28 Prague, Czech Republic
[c] Department of Organic Chemistry, Faculty of Science, Charles University, 116 28 Prague, Czech Republic
[d] Department of Inorganic Chemistry, Faculty of Science, Charles University, 116 28 Prague, Czech Republic

ARTICLE INFO

ABSTRACT

Conventional Nuclear Magnetic Resonance (NMR) analysis relies on H-H/C-H interactions. However, these interactions are sometimes insufficient for an accurate and precise NMR analysis. In this study, we show that ^{31}P NMR parameters can provide critical structural insights into the stereochemistry of phosphorus-containing compounds. For this purpose, we prepared a set of model phosphorus-based proline derivatives, separated diastereoisomers, and determined their absolute configuration by single-crystal X-ray diffraction. After supplementing these results by electronic circular dichroism (ECD) spectroscopy, we combined experimental data and DFT calculations from our model compounds to perform a detailed conformational analysis, thereby determining their relative configuration. Overall, our findings establish an experimental paradigm for combining ^{31}P NMR spectroscopy with other optical methods to facilitate the stereochemical analysis of phosphorus-containing compounds.

© 2022 Elsevier Inc. All rights reserved.

1. Introduction

Phosphorus plays a crucial role in nature as a component of nucleotides, building blocks of nucleic acids (DNA and RNA), and cell membranes (phospholipids) [1]. In addition, phosphorus-based systems enable energy storage (ATP/ADP cycle) and signal transduction [2]. Unsurprisingly, phosphorus is also an essential element in organic chemistry (e.g., organocatalysis) [3], and compounds with a stereogenic center on the phosphorus atom, known as P-chirogenic molecules [4], are used in both enantioselective catalysis and coordination chemistry. In short, whether natural or synthetic, phosphorus compounds support a wide range of functions, with both chemical and medicinal applications.

In medicinal chemistry, P-containing compounds have been used to prepare lipophosphonoxins with antibacterial effects [5] and nucleoside prodrugs with antiviral effects. Among the latter, ProTides [6] stand out as highly potent antiviral drugs against Human Immunodeficiency Virus (HIV), Hepatitis C Virus (HCV), and Severe Acute Respiratory Syndrome Coronavirus-2 (SARS-Cov-2). Another class of P-containing compounds termed phosphorus-based self-immolative (SI) linkers can serve as efficient cargo releasers [7], with advanced phosphorus-based linkers offering multiple modes of double-cargo release [8], e.g., phenols and amines [9].

Structural analysis of organic compounds usually relies on classical analytical methods, such as NMR or optical spectroscopy and mass spectrometry (MS). In turn, stereochemical questions are often addressed by X-ray diffraction data analysis. However, obtaining single crystals of P-containing compounds suitable for X-ray analysis can be a challenging task. For this reason, new methods must be developed for the stereochemical analysis of P-compounds. A possible approach consists of incorporating a magnetically active ^{31}P nucleus and its interaction with neighbouring nuclei into standard NMR analysis when conventional C-H NMR fails.

The NMR active ^{31}P nucleus has a spin ½, a large chemical shift dispersion (1000 ppm), a high natural abundance, and a sufficiently high Larmor frequency and thus does not require a low-gamma NMR probe. Concurrently, proton decoupling affords single ^{31}P NMR signals for each phosphorus component in a reaction mixture. Thanks to these features, ^{31}P spectroscopy can be easily used for reaction monitoring. Another advantage of this approach is that the ^{31}P chemical shift responds to minor changes (e.g., solvation) [10], which may be otherwise confusing when comparing different samples. Spin-spin interactions of ^{31}P with neighbouring nuclei (^{1}H, ^{13}C) provide key connectivity information when screening newly-designed ProTide analogs for drug release properties [11].

Moreover, both ^{13}C–^{31}P and ^{13}C–^{13}C spin–spin interactions have already been used for conformational analysis based on the Karplus-type relationship between J-couplings and dihedral angles [12]. The IGLO-III basis set [13] and the GIAO [14] method were developed for calculating ^{31}P NMR parameters. This method provides computed NMR parameters with reliable accuracy at reasonable computational costs [15].

In this study, we demonstrate that ^{31}P NMR parameters can provide structural insights into the stereochemistry of phosphorus-containing compounds. For this purpose, we prepared a set of model phosphorus-based proline derivatives (Scheme 1). After separating diastereoisomers and determining their absolute configuration by single-crystal X-ray diffraction, we supplemented the results with electronic circular dichroism (ECD) spectroscopy data. Subsequently, we combined advanced NMR analysis with quantum-chemical calculations to explore the conformational space and to determine the relative configuration of our model phosphorus-based proline derivatives.

2. Results and discussion

2.1. Synthesis

Our model compounds were synthesized by cyclisation of (R)-(−) or (S)-(+) −2-pyrrolidinemethanols with the corresponding phosphorodichloridate [16] under CHCl$_3$ reflux (Scheme 1). This reaction proceeded relatively quickly, and no starting material was detected by TLC or NMR spectroscopy after 30 min. We observed that the reaction stereospecificity was considerably affected under anhydrous conditions, leading to 1 and 2 in 84 and 82% de, respectively. In 3, the stereospecificity remained low, even under anhydrous conditions. The resolution of isomers was quite demanding, but we were ultimately able to separate them by multiple flash chromatography on Silica (n-hexane/EtOAc). The isomers of 3 were difficult to separate, requiring multiple flash chromatography on Silica followed by reverse-phase chromatography on C18 (water/acetonitrile). Unfortunately, this separation caused partial decomposition of 3 by opening the oxazolidine ring. Compounds 1 and 2 were isolated as white solids, and two isomers (1-SR and 2-RR), one from each pair of enantiomers, were crystalised from n-hexane/EtOAc (1/1 v/v). All four isomers were characterised by NMR spectroscopy. Based on the synthesis, we used

the first stereochemical descriptor for the proline part and the second for phosphorus. Therefore, e.g., 1-SR defines the stereochemistry on proline as (S) and the stereochemistry on phosphorus as (R).

In addition, 3 was isolated as a colorless oil, but even after many experiments with various solvent systems, we were unable to crystallise this compound and hence to determine its absolute configuration.

2.2. Analysis of compounds 1 and 2: Absolute configuration

We aimed to determine the absolute configuration of the model compounds by X-ray diffraction. Fortunately, compounds 1-SR and 2-RR were successfully crystallised from the mixture of n-hexane/EtOAc (1/1 v/v, 20 mg of the compound in 20 mL). X-ray analysis allowed us to determine the absolute configuration of these two compounds as (S, R) for 1-SR and (R, R) for 2-RR. In both cases, Fig. 1 shows an (R) configuration on the phosphorus atom, with the diastereoisomers differing in the configuration on the proline part (C2), which dictates the shape of the bicyclic proline moiety.

To collect more data on the stereochemistry of non-crystallising diastereomers, we used electronic circular dichroism (ECD) spectroscopy. UV/Vis absorption spectra of 1 and 2 in the UV spectral region (180–300 nm, Figure S39 in the SI) showed spectral bands at 188 nm, 218 nm, and a low-intensity spectral band around 265 nm with a vibronic structure typical for benzene and its derivatives [17]. As expected, ECD spectra of compounds 1 and 2 (Fig. 2) were characterised with spectral bands at ~183 nm, ~195 nm, ~220, and ~260 nm with a vibronic structure and with a sign and intensity reflecting the configuration of discrete compounds [17b]. Thus, we observed the opposite ECD spectra of the enantiomeric pairs 1-SS, 2-RR, and 1-SR, 2-RS with spectral bands at ~183 nm ((−) for 1-SS, 2-RS), ~195 nm ((+) for 1-SS, 2-RS), ~220 nm ((+) for 1-SS, 2-RS), ~260 nm with vibronic bands at 253, 259, 265 and 272 nm ((−) for 1-SS, 2-RS).

2.3. Analysis of compounds 1 and 2: Conformational analysis based on J-couplings

We explored the conformational space of compounds 1 and 2. An ensemble of possible structures was generated by Schrödinger (Maestro/MacroModel) and provided six structures for each of
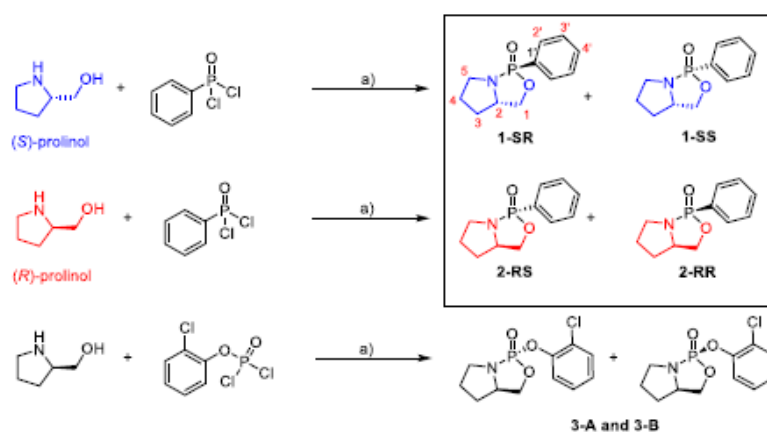
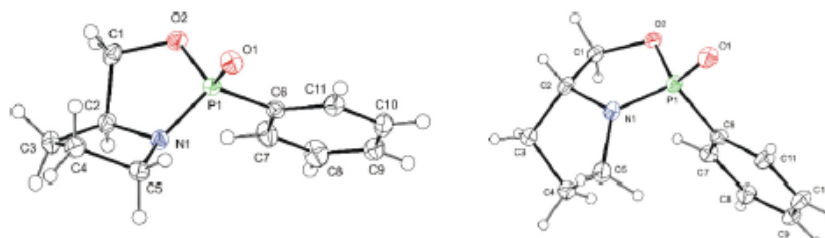Scheme 1. Synthesis of model compounds 1–3 and their atom numbering: a) Et$_3$N (2.2 eq.), CHCl$_3$, reflux 30 min.

Fig. 1. X-ray structures of 1-SR (left) and 2-RR (right); the displacement ellipsoids are drawn on 50% probability level.
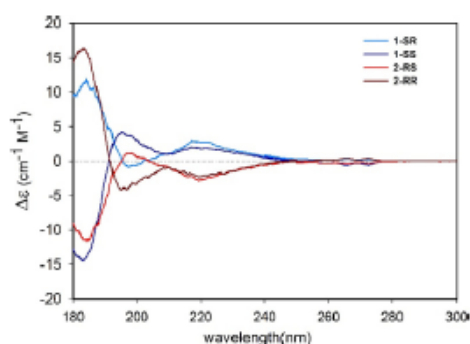

Fig. 2. ECD spectra of compounds 1 and 2 measured in trifluoroethanol at room temperature.

the two isomers of **1** and **2**. These structures were then optimised by density functional theory (DFT) modeling, resulting in three non-redundant conformers as indicated by a careful analysis of dihedral angles.

To exemplify our approach, we show three conformers of **1-SR** and **2-RR**, differing in their relative configuration and thus in the conformation of the proline part. In general, five-membered rings are flexible and can co-exist in several conformations in solution. The stereochemistry of the five-membered ring is usually analysed using the concept of pseudorotation, which was first introduced in cyclopentane by Kilpatrick et al. [18] and later widened by Altona and Sundaralingam [19]. In the theory of pseudorotation, the conformation of the ribose ring is described by two parameters termed phase angle P and maximum puckering amplitude ϕ_{MAX}. By defining the pseudorotation cycle, the furanose ring conformation was schematically represented based on the variation of P (0–360°) [20] (Figures S1 and S2 in the ESI). The cycle is divided into twenty increments (18° step) corresponding to twenty possible conformations of the furanose ring. North (N-type) conformers have

Table 1
Calculated J-couplings (B3LYP/IGLO-III, GD3, PCM = chloroform) of the three conformers and comparison with experimental values extracted from ¹H and ¹³C APT spectra of 1-SR measured in CDCl₃.

Interaction	Calculated J-couplings (Hz)			Boltzmann distribution-weighted average of calculated J-couplings (Hz)	Experimental J-couplings (Hz)
	conf A	conf B	conf C		
H1a-H2	6.3	6.1	5.8	6.2	6.6
H1b-H2	10.0	9.4	0.3	9.3	8.8
H3a-H2	8.0	5.4	4.4	7.3	–
H3b-H2	0.5	9.3	10.8	2.7	–
H5a-H4a	7.1	5.8	8.8	7.0	–
H5a-H4b	0.3	11.7	7.5	2.7	–
H5b-H4a	11.4	0.3	2.1	8.9	–
H5b-H4b	4.5	7.8	10.2	5.4	–
H5a-P	12.6	12.5	17.7	12.9	–
H5b-P	20.7	1.5	17.7	17.2	–
H1a-P	30.5	28.0	−0.1	28.2	20.3
H1b-P	−0.8	−0.7	20.6	0.5	2.4
C1′-P	170.9	170.1	166.0	170.5	182.5
C2′a-P	11.5	11.4	12.0	11.5	10.1
C3′a-P	15.0	14.8	14.8	15.0	15.3
C4′-P	−3.0	−2.9	−3.0	−3.0	3.0
C3′b-P	16.0	16.0	15.6	16.0	15.3
C2′b-P	14.8	15.0	13.3	14.7	10.1
C1-P	3.9	1.1	0.5	3.2	2.3
C2-P	5.5	9.0	7.6	6.2	7.7
C3-P	3.5	3.7	−1.1	3.3	3.1
C4-P	−0.4	7.3	−0.2	0.9	2.3
C5-P	0.8	1.6	−0.8	0.8	1.6
C2-H2	146.7	143.3	143.2	145.9	148.6
C4′-H4′	157.7	157.6	157.7	157.7	161.1

* Experimentally obtained J-couplings were extracted from ¹H and ¹³C APT spectra measured with FID resolution 0.3 and 1.1 Hz, respectively. Based on our estimates, the errors of computed J-couplings were lower than 1 Hz, as proposed in the literature [21].

$P = 0–36°$ and $\phi_0 > 0$ ($^3T_2$, ^3E and $_4T^3$), while South (S-type) conformers have $P = 144–180°$ and ϕ_0 less than 0 ($_3T^2$, ^2E and $^2T_1$; further details are provided in Section 1.1, in ESI). We assumed that the proline ring of **1** and **2** showed similar behavior to that of furanose.

The lowest-energy conformer of **1-SR** (conformer A) is in the South conformation and co-exists with two conformers (B and C) in the North conformation. According to Boltzmann distribution, the population of conformers at 25 °C was identified as 77, 17, and 6% (conformers A, B, and C, respectively). Based on our X-ray analysis, we determined that the South conformer had the same structure as the lowest-energy conformer A found in the conformational search. NMR parameters were averaged based on Boltzmann distribution analysis (Table 1).

The same procedure was applied to the **2-RR** isomer, yielding three conformers. The North conformer showed the lowest energy (Figure S2 and Table S1 in the ESI), matching the structure found by X-ray analysis. Boltzmann population distribution showed 47, 36, and 17% of conformers A, B, and C, respectively. NMR parameters were averaged based on Boltzmann distribution analysis (Table S1).

2.4. Assignment of the relative configuration of 1-SR and 2-RR

First, we assessed whether chemical shifts may be used to determine relative configuration and to distinguish diastereoisomers **1-SR** and **2-RR**. After correlating experimental ^{13}C chemical shifts of **1-SR** and **2-RR** and shielding constants calculated for each conformer, A, B, and C, we averaged the shielding constants based on Boltzmann distributions (e.g., for **2-RR**: conformer A:B:C = 47:36:17) (Figures S4–S7 in the ESI). However, no significant difference was found; therefore, we were unable to distinguish diastereomers using this approach.

Considering that ^{31}P J-couplings are highly sensitive to specific conformers, we carefully analysed J-couplings to determine the relative configuration of **1-SR** and **2-RR**. For this purpose, we compared experimental and calculated J-couplings of both diastereoisomers (Tables S1 and S2 in the ESI). The experimental values were extracted from ^{13}C APT NMR spectra of **1-SR** and **2-RR** measured in chloroform-d. The calculated values were obtained using B3LYP/IGLO-III with empirical dispersion correction GD3 and solvent model PCM = chloroform (details in the Experimental Section). Due to the conformational equilibria, we used weighted-averages of the calculated values based on Boltzmann distribution at 25 °C. Ultimately, ^{31}P interactions enabled us to distinguish the diastereoisomers (Fig. 3).

2.5. RDC analysis of 1 and 2

Residual dipolar couplings (RDCs) were extracted by comparing data measured under isotropic (in CDCl$_3$) and anisotropic (in poly-γ-benzyl-L-glutamate (PBLG) and CDCl$_3$) conditions. We entered these experimental data, together with the conformers determined by conformational analysis, into the fitting software.

We used three methods based on different principles: MSpin [22], RDC@hotfcht [23], and P3D/PALES [24]. To exemplify our approach, we show the analysis of **1-SR** in detail; the other diastereomers were treated similarly.

The conformational analysis of **1-SR** provided three conformers (A, B, and C). First, the conformers combined with the experimentally determined RDCs were entered into MSpin software, one of the first commercially available softwares for RDC analysis. All three conformers had good quality factors Q and Pearson's correlation factors R: $Q_{confA} = 0.0133$, $Q_{confB} = 0.0123$, $Q_{confC} = 0.0194$ and $R_{confA} = 0.9999$, $R_{confB} = 0.9998$, $R_{confC} = 0.9997$, respectively (Figure S14 in the ESI). The same experimental RDC data of **1-SR** were then correlated with the lowest-energy conformers A of the other three diastereomers to assess the ability of the method to discriminate diastereoisomers. However, the results did not differ significantly. As a result, we could not define the correct diastereomer using MSpin (Figure S16 in ESI).

The RDC@hotfcht approach is based on a similar principle to that of MSpin, but RDC@hotfcht also includes the experimental error in the fitting procedure. Nevertheless, the results were similar to those of MSpin - the RDC analysis of the conformers yielded generally good results for all three conformers of **1-SR**. Yet again, however, we were unable to differentiate the diastereoisomers **1-SR** and **2-RR** ($Q_{RR} = 0.0202$, $R_{RR} = 0.9998$; $Q_{RS} = 0.0136$, $R_{RS} = 0.9999$; $Q_{SR} = 0.0136$, $R_{SR} = 0.9999$; $Q_{SR} = 0.0202$, $R_{SR} = 0.9998$) (Figures S13 and S15 in the SI).

Searching for a more realistic approach considering the nature of the alignment medium, we used the recently developed software P3D implemented in the PALES program. P3D/PALES simulates the interaction of the analyte with the alignment medium. This method was designed for poly-γ-benzyl-L-glutamate (PBLG) in chloroform, which we used in our study. Moreover, the P3D method was designed to discriminate diastereoisomers. In contrast to MSpin and RDC@hotfcht, where Q-factor is used to evaluate the goodness of fit, the Pearson's correlation factor R is used in P3D/PALES. According to the literature [24], R ≥ 0.8 is a satisfactory result. In our results, the R values of the individual conformers of **1-SR** were $R_{confA} = 0.785$, $R_{confB} = 0.692$, and $R_{confC} = 0.655$ (Figure S12 in the ESI). As such, the lowest-energy conformer A was the most promising for a diastereomer discrimination analysis.

This discrimination analysis, however, identified the correct diastereomer as (R, R) (or (S, S) as enantiomers cannot be discriminated using this method) (Fig. 4). This result did not match the (S, R) configuration determined by X-ray diffraction. Consequently, this method did not discriminate **1-SR**, **1-SS**, and **2-RR**, **2-RS** diastereoisomeric pairs.

The results of the other three isomers are summarized in Tables S6 and S7 in the ESI. Due to the similar results from the MSpin and RDC@hotfcht softwares, we decided to continue using only RDC@hotfcht because this software considers also the experimental errors.

2.6. Analysis of compound 3

Based on the results from model compounds **1** and **2**, we prepared an ortho-chlorophenyl derivative of **2**, which should offer a higher number of ^{13}C-^1H J-couplings and corresponding RDCs due to the disturbed symmetry of the phenyl ring. However, we were unable to prepare such a chlorophenyl analogue. Therefore, we prepared the phosphate derivative **3** (Scheme 1), which is more flexible than **1** and **2**. The addition of the oxygen atom between the phosphorus and the phenyl ring in **3** also prevented a direct comparison between the configurations of **1** and **2** due to the change in substituent priorities according to Cahn-Ingold-Prelog rules [25], as shown in Fig. 5.

Fig. 3. Experimental (pink) and calculated (black) ^{13}C,^{31}P J-couplings of diastereoisomers 1-SR and 2-RR enabled us to determine their relative configuration. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
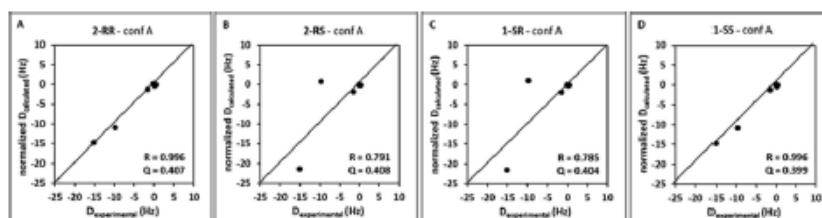
Fig. 4. Correlation of experimental RDCs of **1-SR** and normalised RDC values calculated using P3D/PALES fitted to conformer A of: A – **2-RR**, B – **2-RS**, C – **1-SR**, and D – **1-SS** structures.

Compound **3** did not crystalise. For this reason, we could not perform X-ray diffraction analysis and instead measured ECD spectra of both diastereoisomers. The comparison of the ECD spectra of **3** with **1** and **2** showed that **3-A** has the same geometry on phosphorus as **1-SR** and **2-RR**, while **3-B** has the same geometry as **1-SS** and **2-RS**. Because **3-A** has the same geometry as **2-RR**, **3-A** should correspond to the (R,S) isomer, as shown in Fig. 5. However, in **3-A** and **3-B**, the phenyl is not close to the stereocenter, which further complicated the absolute configuration assignment (details in ESI). ECD spectroscopy should be complemented with methods of vibrational (vibrational circular dichroism (VCD)) and Raman (ROA) optical activity and combined with quantum-chemical calculations [26]. This approach is currently being applied to our systems.

2.7. Analysis of compound 3: Conformational analysis

Notwithstanding these challenges, we searched the conformational space of **3**. Our conformational analysis identified more conformers of **3** than in **1** and **2** due to the higher molecular flexibility of **3** (Tables S3 and S5 in ESI). The conformation of the bicyclic part is dictated by the configuration of prolinol ((R) configuration on C2). The lowest-energy conformer A of **3-RS** and **3-RR** isomer was the North conformation with a phase angle P of 25° and 18°, respectively. In both isomers, the North conformation was in equilibrium with the South-conformation, with P of 0° and –18° for the **3-RS** and **3-RR**, respectively. Ultimately, we identified nine conformers of the **3-RS** isomer and seventeen of the **3-RR** isomer. Population analysis of **3** based on Boltzmann distribution showed that **3-RS** had five (A–E) and **3-RR** nine (A–I) conformers; the other eight conformers accounted for less than 2% of the total population and were not used for averaging of the NMR parameters (shielding constants, J-couplings, details in Section 1, in ESI).

The lowest-energy conformer A of **3-RS** had the phenoxy group in an equatorial geometry with a dihedral angle N-P-O-C$_{ipso}$ = 176°.

In contrast, the conformer A of **3-RR** isomer was in an axial position, with a dihedral angle N-P-O-C$_{ipso}$ = –39°.

We calculated the J-couplings of the selected conformers of **3**, but the differences in this parameter were not significant enough to discriminate the diastereomers (Table S4 in ESI). The correlations between the chemical shifts of **3-A** and **3-B** and the calculated shielding constants of both diastereomers **3-RS** and **3-RR** did not allow us to differentiate the diastereomers either (Figures S8–S11, in ESI).

2.8. Analysis of compound 3: RDC analysis

The RDC analysis of **3** was performed as described for **1** and **2**. The Cl atom attached to the phenyl ring broke down the symmetry and enabled us to obtain more one-bond ^{13}C–^{1}H RDCs from the phenyl moiety. Consequently, two- or three-bond ^{13}C-^{31}P RDCs should improve differentiation between conformers and diastereoisomers. Unfortunately, the RDC@hotfcht software did not indicate the correct structure of **3-A** because almost all the resulting Pearson's correlation coefficients R of both **3-RR** and **3-RS** isomers were higher than 0.9 (Figures S32 and S36 in the ESI). The RDC@hotfcht results of **3-B** were similar and did not distinguish the diastereoisomers. Nevertheless, the P3D/PALES software established **3-A** as a **3-RR** isomer (conformer F) with R = 0.857, but **3-B** remained undetermined (the best **3-RR** conformer was M, with R = 0.658). These results indicate that using only the lowest-energy conformers in the RDC analysis may not be the correct approach. All RDC data are summarised in Tables S8 and S9, and Figures S31–S38, in ESI.

3. Conclusion

^{31}P J-couplings strongly depend on structural changes. This correlation may be useful for constructing Karplus-type curves within structurally diverse classes of compounds. The RDC couplings pro-
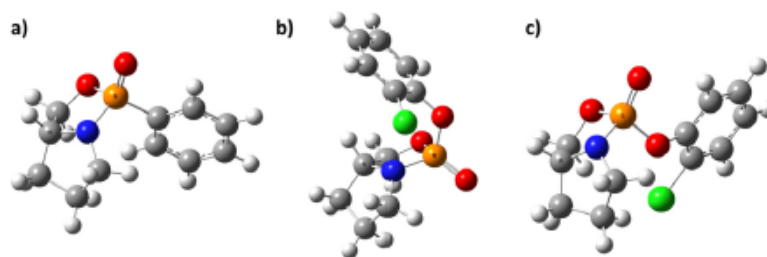

Fig. 5. Comparison of the calculated lowest-energy conformers A of a) **2-RR**, b) **3-RR**, and c) **3-RS**.

vided ambiguous results possibly because the absolute values of the two- and three-bond ^{13}C-^{31}P RDC couplings were low (approximately one order of magnitude smaller than one-bond couplings). As a result, the RDCs were very small and did not significantly improve the fitting procedure. RDC@hotfcht provided similar results for all studied compounds, **1**, **2**, and **3**, whose Pearson's correlation coefficients R were all higher than 0.9. P3D/PALES, a new program designed for diastereoisomer discrimination, improved the results. However, we were unable to differentiate the diastereomers unambiguously. Additional problems have arisen during the analysis of the more flexible derivative **3**. An alternative option may be to use of the MDOC approach [27], albeit necessarily modified for ^{31}P interactions. Our results suggest that ^{31}P J-couplings combined with ^{31}P NMR parameters calculations may facilitate the stereochemical analysis of small molecules containing phosphorus, but a viable process has not been developed yet given the aforementioned difficulties.

4. Experimental Section

4.1. Synthesis

General. All reagents were purchased from commercial suppliers and used as received. Thin-layer chromatography (TLC) was performed on TLC aluminium sheets (silica-gel 60 F254; Merck). Reaction progress was monitored by TLC and/or ^{31}P NMR spectroscopy in CDCl$_3$. Flash-column chromatography was performed on a Compact (ECOM s.r.o.) chromatography system using silica-gel, 100 g, 230–400 mesh, 60 Å (Merck) in a gradient mode (n-hexane/EtOAc, 0–100%, 40 mL/min, v/v) or Prontosil HPLC column RP C18 (20 mm, 250 mm) also in a gradient mode (water/acetonitrile, 20–80%, 25 mL/min, v/v). All products were obtained as white solids or colorless oils. The reaction yields were not optimised.

(1R, 3aS)-1-Phenyltetrahydro-3H-pyrrolo[1,2-c][1,3,2]oxaza phosphole 1-oxide. (1-SR). (S)-(+)-2-Pyrrolidinmethanol (990 μl, 10 mmol, 1 equiv.) and triethylamine (3 mL, 22 mmol, 2.2 equiv.) were dissolved in dry chloroform (20 mL) under argon atmosphere. Phenylphosphonic dichloride (1.4 mL, 10 mmol, 1 equiv.) was subsequently added in dry chloroform (10 mL) under reflux. The reaction mixture was further refluxed under argon atmosphere for 30 min. Reaction conversion was monitored by TLC. The reaction mixture was extracted with water, and the organic layer was dried with MgSO$_4$ and evaporated to dryness *in vacuo*. The reaction proceeded with a diastereoisomeric ratio of 92:8 (SR:SS), 84% de. The title diastereoisomer was purified by flash chromatography on a silica gel using a gradient (n-hexane/EtOAc, 0–100%, v/v). The compound was isolated as a white solid (1.3 g, 58%).

^{1}H NMR (400 MHz, CDCl$_3$, 25 °C): δ = 7.81–7.88 (m, 2H, 2'), 7.54 (m, 1H, 4'), 7.43–7.49 (m, 2H, 3'), 4.34 (ddd, 1H, J$_{1a-P}$ = 20.3, J$_{GEM}$ = 8.8, J$_{1a-2}$ = 6.6, 1a), 4.15 (m, 1H, 2), 3.92 (dm, 1H, J$_{1b-P}$ = 2.4, 1b), 3.77 (m, 1H, 5a), 2.95 (m, 1H, 5b), 1.99–2.11 (m, 3H, 3a, 4), 1.82 ppm (m, 1H, 3b); ^{13}C NMR (100 MHz, CDCl$_3$, 25 °C): δ = 131.98 (d, J$_{4'-P}$ = 3.0, 4'), 131.62 (d, J$_{2'-P}$ = 10.1, 2'), 130.88 (d, J$_{1'-P}$ = 182.5, 1'), 128.32 (d, J$_{3'-P}$ = 15.3, 3'), 69.67 (d, J$_{5-P}$ = 2.3, 1), 63.20 (d, J$_{2-P}$ = 7.7, 2), 45.35 (d, J$_{5-P}$ = 1.6, 5), 29.93 (d, J$_{3-P}$ = 3.1, 3), 27.53 ppm (d, J$_{4-P}$ = 2.3, 4); ^{31}P NMR (161 MHz, CDCl$_3$, 25 °C): δ = 38.67 ppm.

HR-MS (APCI) calculated for C$_{11}$H$_{15}$O$_2$NP 224.08349, found [M + H]$^+$ 224.08342.

[α]$_{589}^{20}$ = 77.2°; 0.378 g/100 mL in chloroform.

Crystal data for **1-SR**: C$_{11}$H$_{14}$NO$_2$P, M_r = 223.20; Orthorhombic, P 2$_1$ 2$_1$ 2$_1$, (No 19), a = 6.4205 (2) Å, b = 9.0993 (3) Å, c = 18.6354 (5), V = 1088.72 (6) Å^3, Z = 4, D_x = 1.362 Mg m^{-3}, Prism, colorless crystal of dimensions 0.25 × 0.20 × 0.14 mm, multi-scan absorption correction (μ = 2.08 mm^{-1}) T_{min} = 0.67, T_{max} = 0.76; a total of 17,053 measured reflections (θ$_{max}$ = 77.3°), from which 2261 were unique (R_{int} = 0.023) and 2254 observed according to the $I > 2\sigma(I)$ criterion.

The refinement converged (Δ/σ$_{max}$ = 0.001) to R = 0.023 for observed reflections and wR(F^2) = 0.062, GOF = 1.07 for 136 parameters and all 2261 reflections. The final difference density map displayed no peaks of chemical relevance (Δρ$_{max}$ 0.22, Δρ$_{min}$ −0.19 e.Å^{-3}). Absolute structure parameter: −0.002 (4) [28].

(1S, 3aS)-1-Phenyltetrahydro-3H-pyrrolo[1,2-c][1,3,2]oxaza phosphole 1-oxide. (1-SS). (S)-(+)-2-Pyrrolidinmethanol (990 μl, 10 mmol, 1 equiv.) and triethylamine (3 mL, 22 mmol, 2.2 equiv.) were dissolved in dry chloroform (20 mL). Phenylphosphonic dichloride (1.4 mL, 10 mmol, 1 equiv.) was subsequently added in dry chloroform (10 mL) under reflux. The reaction mixture was further refluxed without argon atmosphere for 30 min. Reaction conversion was monitored by TLC. The reaction mixture was extracted with water, and the organic layer was dried with MgSO$_4$ and evaporated to dryness *in vacuo*. The reaction proceeded with a diastereoisomeric ratio of 67:33 (SR:SS), 34% de. The title diastereoisomer was purified by flash chromatography on silica gel using a gradient (n-hexane/EtOAc, 0–100%, v/v). The compound was isolated as a white solid (250 mg, 11%).

^{1}H NMR (400 MHz, CDCl$_3$, 25 °C): δ = 7.74–7.81 (m, 2H, 2'), 7.56–7.62 (m, 1H, 4'), 7.48–7.54 (m, 2H, 3'), 4.75 (ddd, 1H, J$_{1a-P}$ = 15.4, J$_{GEM}$ = 9.0, J$_{1a-2}$ = 6.4, 1a), 4.23–4.32 (m, 1H, 2), 4.06 (ddd, J$_{1b-P}$ = 13.3, J$_{GEM}$ = 9.0, J$_{1b-2}$ = 4.3, 1b), 3.01–3.10 (m, 1H, 5a), 2.84–2.92 (m, 1H, 5b), 2.02–2.12 (m, 1H, 3a), 1.90–2.01 (m, 1H, 4a), 1.78–1.89 (m, 1H, 4b), 1.55–1.66 ppm (m, 1H, 3b); ^{13}C NMR (100 MHz, CDCl$_3$, 25 °C): δ = 132.33–132.51 (m, 2', 4'), 128.70 (d, J$_{3'-P}$ = 13.9, 3'), 127.6** (1'), 74.16 (d, J$_{1-P}$ = 1.6, 1), 61.49 (d, J$_{2-P}$ = 13.1, 2), 44.55 (d, J$_{5-P}$ = 6.2, 5), 30.40 (d, J$_{3-P}$ = 4.7, 3), 27.44 ppm (d, J$_{4-P}$ = 4.6, 4); ^{31}P (161 MHz, CDCl$_3$, 25 °C): δ = 34.05 ppm.

HR-MS (APCI) calculated for C$_{11}$H$_{15}$O$_2$NP 224.08349, found [M + H]$^+$ 224.08340.

[α]$_{589}^{20}$ = −15.6°; 0.199 g/100 mL in chloroform.

**extracted from HMBC

(1S, 3aR)-1-Phenyltetrahydro-3H-pyrrolo[1,2-c][1,3,2]oxaza phosphole 1-oxide. (2-RS). (R)-(-)-2-Pyrrolidinmethanol (990 μl, 10 mmol, 1 equiv.) and triethylamine (3 mL, 22 mmol, 2.2 equiv.) were dissolved in dry chloroform (20 mL) under argon atmosphere. Phenylphosphonic dichloride (1.4 mL, 10 mmol, 1 equiv.) was subsequently added in dry chloroform (10 mL) under reflux. The reaction mixture was further refluxed under argon atmosphere for 30 min. Reaction conversion was monitored by TLC. The reaction mixture was extracted with water, and the organic layer was dried with MgSO$_4$ and evaporated to dryness *in vacuo*. The reaction proceeded with a diastereoisomeric ratio of 91:9 (RS:RR), 82% de. The title diastereoisomer was purified by flash chromatography on silica gel using a gradient (n-hexane/EtOAc, 0–100%, v/v). The compound was isolated as a white solid (1 g, 45%).

^{1}H NMR (400 MHz, CDCl$_3$, 25 °C): δ = 7.80–7.88 (m, 2H, 2'), 7.51–7.57 (m, 1H, 4'), 7.43–7.50 (m, 2H, 3'), 4.35 (ddd, 1H, J$_{1a-P}$ = 15.3, J$_{GEM}$ = 8.8, J$_{1a-2}$ = 6.6, 1a), 4.11–4.20 (m, 1H, 2), 3.92 (ddd, J$_{1b-P}$ = 11.1, J$_{GEM}$ = 8.7, J$_{1b-2}$ = 2.5, 1b), 3.73–3.82 (m, 1H, 5a), 2.89–3.00 (m, 1H, 5b), 1.99–2.11 (m, 3H, 3a, 4), 1.79–1.87 ppm (m, 1H, 3); ^{13}C NMR (100 MHz, CDCl$_3$, 25 °C): δ = 132.00 (d, J$_{4'-P}$ = 3.1, 4'), 131.63 (d, J$_{2'-P}$ = 10.5, 2'), 129.8** (1'), 128.34 (d, J$_{3'-P}$ = 14.6, 3'), 69.69 (d, J$_{1-P}$ = 2.3, 1), 63.22 (d, J$_{2-P}$ = 7.8, 2), 45.36 (5), 29.95 (d, J$_{3-P}$ = 3.1, 3), 27.55 ppm (d, J$_{4-P}$ = 1.5, 4); ^{31}P NMR (161 MHz, CDCl$_3$, 25 °C): δ = 38.73 ppm.

HR-MS (APCI) calculated for C$_{11}$H$_{15}$O$_2$NP 224.08349, found [M + H]$^+$ 224.08346.

[α]$_{589}^{20}$ = −64.5°; 0.380 g/100 mL in chloroform.

**extracted from HMBC

(1R, 3aR)-1-Phenyltetrahydro-3H-pyrrolo[1,2-c][1,3,2]oxaza-phosphole 1-oxide. (2-RR). (R)-(-)-2-Pyrrolidinmethanol (990 μl, 10 mmol, 1 equiv.) and triethylamine (3 mL, 22 mmol, 2.2 equiv.) were dissolved in dry chloroform (20 mL). Phenylphosphonic

dichloride (1.4 mL, 10 mmol, 1 equiv.) was subsequently added in dry chloroform (10 mL) under reflux. The reaction mixture was further refluxed without argon atmosphere for 30 min. Reaction conversion was monitored by TLC. The reaction mixture was extracted with water, organic layer was dried with $MgSO_4$ and evaporated to dryness *in vacuo*. The reaction proceeded with a diastereoisomeric ratio of 67:33 (*RS*:*RR*), 34% de. The title diastereoisomer was purified by flash chromatography on silica gel using a gradient (*n*-hexane/EtOAc, 0–100%, v/v). The compound was isolated as a white solid (300 mg, 13%).

^1H NMR (400 MHz, $CDCl_3$, 25 °C): δ = 7.72–7.80 (m, 2H, 2'), 7.55–7.61 (m, 1H, 4'), 7.47–7.53 (m, 2H, 3'), 4.74 (ddd, 1H, $J_{1a–P}$ = 15.4, J_{GEM} = 9.0, $J_{1a–2}$ = 6.4, 1a), 4.22–4.31 (m, 1H, 2), 4.05 (ddd, $J_{1b–P}$ = 13.2, J_{GEM} = 9.0, $J_{1b–2}$ = 4.3, 1b), 3.00–3.09 (m, 1H, 5a), 2.82–2.91 (m, 1H, 5b), 2.01–2.11 (m, 1H, 3a), 1.89–2.00 (m, 1H, 4a), 1.77–1.88 (m, 1H, 4b), 1.54–1.64 ppm (m, 1H, 3b); ^{13}C NMR (100 MHz, $CDCl_3$, 25 °C): δ = 132.41 (d, $J_{2'–P}$ = 10.8, 2'), 131.34 (m, 4'), 128.70 (d, $J_{3'–P}$ = 14.5, 3'), 127.97 (d, $J_{1'–P}$ = 162.7, 1'), 74.16 (1), 61.49 (d, $J_{2–P}$ = 13.2, 2), 44.55 (d, $J_{5–P}$ = 6.2, 5), 30.41 (d, $J_{3–P}$ = 4.6, 3), 27.44 ppm (d, $J_{4–P}$ = 4.6, 4); ^{31}P (161 MHz, $CDCl_3$, 25 °C): δ = 34.03 ppm.

HR-MS (APCI) calculated for $C_{11}H_{15}O_2NP$ 224.08349, found [M + H]$^+$ 224.08352.

$[\alpha]_{589}^{20}$ = 13.1°; 0.335 g/100 mL in chloroform.

Crystal data for ae42b: $C_{11}H_{14}NO_2P$, M_r = 223.20; Orthorhombic, P 2$_1$ 2$_1$ 2$_1$, (No 19), a = 6.3531 (3)Å, b = 12.8725 (6)Å, c = 13.2265 (6), V = 1081.67 (9) Å^3, Z = 4, D_x = 1.371 Mg m^{-3}, Prism, colorless crystal of dimensions 0.18 × 0.10 × 0.08 mm, multi-scan absorption correction (μ = 2.09 mm^{-1}) T_{min} = 0.70, T_{max} = 0.85; a total of 7405 measured reflections (θ_{max} = 72.1°), from which 2124 were unique (R_{int} = 0.046) and 1961 observed according to the $I > 2\sigma(I)$ criterion. The refinement converged (Δ/σ_{max} = 0.001) to R = 0.032 for observed reflections and wR(F^2) = 0.078, GOF = 1.09 for 136 parameters and all 2124 reflections. The final difference density map displayed no peaks of chemical relevance ($\Delta\rho_{max}$ = 0.21, $\Delta\rho_{min}$ −0.21 e.Å^{-3}). Absolute structure parameter: −0.006 (14) [28].

(1*R* or 1*S**, 3a*R*)-1-(2-chlorophenoxy)tetrahydro-3H-pyrrolo [1,2-c][1,3,2]oxazaphosphole 1-oxide. (3A) (*R*)-(-)-2-Pyrrolidinmethanol (990 μl, 10 mmol, 1 equiv.) and triethylamine (3 mL, 22 mmol, 2.2 equiv.) were dissolved in dry chloroform (20 mL). 2-Chlorophenyl phosphorodichloridate (1.6 mL, 10 mmol, 1 equiv.) was subsequently added in dry chloroform (10 mL) at reflux. The reaction mixture was further refluxed without argon atmosphere for 30 min. Reaction conversion was monitored by TLC. The reaction mixture was extracted with water, organic layer was dried with $MgSO_4$ and evaporated to dryness *in vacuo*. The reaction proceeded with a diastereoisomeric ratio of 25:75, 50% de. The title diastereoisomer was purified by flash chromatography on silica gel using a gradient (*n*-hexane/EtOAc, 0–100%, v/v), followed by C18 using a gradient (water/acetonitrile). The compound was isolated as a colorless oil (475 mg, 18%). *We were unable to determine absolute configuration.

^1H NMR (500 MHz, $CDCl_3$, 25 °C): δ = 7.43 (dm, 1H, $J_{3'–4'}$ = 8.0, 3'), 7.40 (dm, 1H, $J_{6'–5'}$ = 8.2, 6'), 7.25 (ddd, 1H $J_{5'–3'}$ = 1.5, $J_{5'–4'}$ = 7.8, $J_{5'–6'}$ = 9.1, 5'), 7.13 (ddm, 1H, $J_{4'–6'}$ = 1.3, $J_{4'–3}$ = 7.8, 4'), 4.34 (ddd, 1H, $J_{1a–2}$ = 6.6, J_{GEM} = 8.8, $J_{1a–P}$ = 21.8, 1a), 3.92 (m, 1H, 1b), 3.83 (m, 1H, 2), 3.75 (m, 1H, 5a), 3.07 (m, 1H, 5b), 1.92–2.06 (m, 3H, 3a, 4a, 4b), 1.69 ppm (m, 1H, 3b); ^{13}C NMR (125 MHz, $CDCl_3$, 25 °C): δ = 147.00 (d, $J_{1'–P}$ = 7.0, 1'), 130.49 ($J_{3'–P}$ = 1.5, 3'), 127.74 (d, $J_{5'–P}$ = 1.9, 5'), 126.36 (d, $J_{2'–P}$ = 6.2, 2'), 125.81 (d, $J_{4'–P}$ = 2.3, 4'), 122.32 (d, $J_{6'–P}$ = 3.1, 6'), 70.36 (d, $J_{1–P}$ = 4.6, 1), 61.41 (d, $J_{2–P}$ = 12.7, 2), 46.71 (d, $J_{5–P}$ = 3.3, 5), 30.75 (d, $J_{3–P}$ = 3.9, 3), 26.77 ppm (d, $J_{4–P}$ = 3.7, 4); ^{31}P (202 MHz, $CDCl_3$, 25 °C): δ = 21.74 ppm.

HR-MS (APCI) calculated for $C_{11}H_{14}O_3NClP$ 274.03943, found [M + H]$^+$ 274.03946.

$[\alpha]_{589}^{20}$ = −48.1°; 0.339 g/100 mL in chloroform.

(1*S* or 1*R**, 3a*R*)-1-(2-chlorophenoxy)tetrahydro-3H-pyrrolo [1,2-c][1,3,2]oxazaphosphole 1-oxide. (3B). (*R*)-(-)-2-Pyrrolidinmethanol (990 μl, 10 mmol, 1 equiv.) and triethylamine (3 mL, 22 mmol, 2.2 equiv.) were dissolved in dry chloroform (20 mL). 2-Chlorophenyl phosphorodichloridate (1.6 mL, 10 mmol, 1 equiv.) was subsequently added in dry chloroform (10 mL) under reflux. The reaction mixture was further refluxed without argon atmosphere for 30 min. Reaction conversion was monitored by TLC. The reaction mixture was extracted with water, and the organic layer was dried with $MgSO_4$ and evaporated to dryness *in vacuo*. The reaction proceeded with a diastereoisomeric ratio of 29:71, 42% de. The title diastereoisomer was purified by flash chromatography on silica gel using a gradient (*n*-hexane/EtOAc, 0–100%, v/v), followed by C18 using a gradient (water/acetonitrile). The compound was isolated as a colorless oil (1.43 g, 54%). *We were unable to determine absolute configuration.

^1H NMR (400 MHz, $CDCl_3$, 25 °C): δ = 7.47 (dm, 1H, $J_{6'–5'}$ = 8.2, 6'), 7.41 (dm, 1H, $J_{3'–4}$ = 8.0, 3'), 7.26 (ddd, 1H, $J_{5'–4'}$ = 7.5, $J_{5'–3}$ = 1.7, $J_{5'–6}$ = 8.2, 5'), 7.11 (dddd, 1H, $J_{4'–6'}$ = 0.9, $J_{4'–6'}$ = 1.5, $J_{4'–5'}$ = 7.5, $J_{4'–3'}$ = 8.0, 4'), 4.52 (ddd, 1H, $J_{1a–2}$ = 6.6, J_{GEM} = 9.0, $J_{1a–P}$ = 23.2, 1a), 4.26 (m, 1H, 2), 4.01 (td, 1H, $J_{1b–P}$ = 1.0, J_{GEM} = 9.0, $J_{1b–2}$ = 9.0, 1b), 3.74 (m, 1H, 5a), 3.12 (m, 1H, 5b), 2.00–2.20 (m, 3H, 4, 3a), 1.72 ppm (m, 1H, 3b); ^{13}C NMR (100 MHz, $CDCl_3$, 25 °C): δ = 147.40 (d, $J_{1'–P}$ = 8.9, 1'), 130.36 (3'), 127.96 (d, $J_{5'–P}$ = 1.6, 5'), 125.53 (4'), 125.20 (d, $J_{2'–P}$ = 7.7, 2'), 121.79 (6'), 71.45 (d, $J_{1–P}$ = 4.4, 1), 62.26 (d, $J_{2–P}$ = 13.1, 2), 45.36 (d, $J_{5–P}$ = 1.6, 5), 29.94 (d, $J_{3–P}$ = 3.9, 3), 27.66 ppm (d, $J_{4–P}$ = 3.1, 4); ^{31}P (161 MHz, $CDCl_3$, 25 °C): δ = 17.01 ppm.

HR-MS (APCI) calculated for $C_{11}H_{14}O_3NClP$ 274.03943, found [M + H]$^+$ 274.03938.

$[\alpha]_{589}^{20}$ = 11.7°; 0.326 g/100 mL in chloroform.

4.2. X-ray diffraction

The X-ray experiments for structure determination were performed on a Bruker D8 VENTURE single-crystal Kappa-axis diffractometer with a Duo PHOTONIII detector and a IμS micro-focus sealed tube CuKα (λ = 1.54178 Å) at a temperature of 120(2) K. The structure was solved using direct methods (XT [29]) and refined by full matrix least squares based on F^2 (SHELXL2018 [30]). The hydrogen atoms on carbon were fixed into idealised positions (riding model) and assigned temperature factors $H_{iso}(H)$ = 1.2 U$_{eq}$(pivot atom). The absolute configuration assignment was based on anomalous scattering of P, O and N atoms.

X-ray crystallographic data were deposited with the Cambridge Crystallographic Data Centre (CCDC) under deposition number 2122002, 2122001 for **1-SR** and **2-RR**, respectively, and can be obtained free of charge from the website of the Centre at www.ccdc.cam.ac.uk/getstructures.

4.3. Circular dichroism spectroscopy

ECD spectra were measured on a Jasco 815 spectropolarimeter over a spectral range of 180 nm to 300 nm in trifluoroethanol (TFE) in a quartz cell with a 0.02 cm (180–250 nm) and 0.1 cm (230–300 nm) path length using a scanning speed of 10 nm/min, a response time of 8 s and 5 accumulations with standard instrument sensitivity at a concentration of 2.0 × 10^{-4} M and 1.0 × 10^{-3} M, respectively. After baseline correction, spectra were expressed in terms of differential molar extinction ($\Delta\varepsilon$).

4.4. Optical rotation

Optical rotations were measured in chloroform using an Autopol IV instrument (Rudolph Research Analytical).

4.5. Conformational search

The conformational sampling implemented in MacroModel v13.1 (Schrödinger 2021-1 suite) [31] was performed using the following set of parameters: force field OPLS4 [32], solvent chloroform, mixed torsional/low-mode sampling method with a maximum of 1000 steps, an energy window for saving structures of 40 kJ/mol, and a maximum atom deviation cut-off of 0.75 Å.

4.6. Pseudorotation cycle analysis

Dihedral angles ϕ_0–ϕ_4 were extracted from DFT-optimised structures, and the phase angle P determining the shape of the proline five-membered ring was calculated from the following equation:

$$\tan P = \frac{(\Phi_2 + \phi_4) - (\Phi_1 + \phi_3)}{2\phi_0(sin36° + sin72°)}$$

4.7. DFT calculations

DFT calculations were performed using the Gaussian 16 package [33]. Geometry optimisations of the conformers obtained from conformational sampling were conducted employing the B3LYP functional [34] with a 6-31+G(d,p) [35] basis set with empirical dispersion correction (GD3) [36]. Calculations were performed *in vacuo* and using the polarisable continuum model (PCM [37] chloroform). The NMR parameters were calculated using the B3LYP functional with a basis set optimised for ^{31}P parameters (IGLO-III) [12b, 15, 21].

4.8. Sample preparation

The quantity of the alignment media was derived from the quantities of the solvent and analyte to acquire approximately 7.7 wt%. This number was determined based on our previous experience and on a premeasurement screening. The calculated amount of the alignment medium was mixed with the respective amounts of the solvent, and the analyte and the final mixture was left standing to dissolve overnight. Due to the high viscosity of the solution, the sample was homogenised using a manual centrifuge. The sample homogeneity was then monitored by ^2H image experiments, and the alignment order was determined by measuring the quadrupolar splitting of the solvent in ^2H NMR.

Extraction of experimental RDCs and their evaluation

The residual dipolar couplings were acquired from F1-coupled HSQC experiments (one-bond ^{13}C-^1H couplings) or APT experiments (^{13}C-^{31}P couplings) using the following equation:

$$^nD_{C-Y} = \frac{^nT_{C-Y} - ^nJ_{C-Y}}{2},$$

where J stands for the scalar coupling (obtained from an isotropic experiment - measured in $CDCl_3$), T is the total coupling (collected from an anisotropic experiment measured in the solvent $CDCl_3$ and in the alignment medium PBLG), D is the residual dipolar coupling, index ^{n} indicates the number of bonds and C-Y stands for ^{13}C-^1H or ^{13}C-^{31}P coupling. The RDC data were evaluated using Pearson's correlation coefficient R and Quality factor Q.

The Pearson's correlation coefficient R is defined as follows:

$$R = \frac{n\sum D_{exp}D_{calc} - \sum D_{exp}\sum D_{calc}}{\sqrt{n\sum D_{exp}^2 - (\sum D_{exp})^2}\sqrt{n\sum D_{calc}^2 - (\sum D_{calc})^2}}$$

where n is the number of experimental RDC values, D_{exp} are the experimental RDCs and D_{calc} are the calculated RDC values (from an appropriate software).

The quality factor is given by the following equation:

$$Q = \sqrt{\frac{\frac{1}{n}\sum(D_{exp}^2 - D_{calc}^2)}{\frac{1}{n}\sum D_{exp}^2}},$$

where D_{exp} are the experimental RDC values, D_{calc} are the calculated RDCs, and n is the number of the experimental RDCs.

4.9. RDC analysis using RDC@hotFCHT [23]

The RDC calculations using the RDC module implemented in the hotFCHT software were executed in accordance with the standard procedure. The input file contained experimental RDCs and xyz coordinates of the optimised geometry of the analyte in a format suitable for RDC@hotFCHT. The xyz coordinates of the optimised structure were obtained as mentioned above in the section *DFT calculations*.

4.10. RDC analysis using MSpin [22]

MSpin software with the RDC module performs the least-square determination of the alignment tensor using a singular value decomposition computational algorithm.

4.11. RDC analysis using P3D/Pales

Molecular alignment simulations using P3D [24a] as implemented in the PALES [24b] software were performed as recommended [38]. The following command was used to run the simulation:

pales -elPales -3D -pot3D PBLG.dx -lcS 0.8 -maxPot 2 -z1 150 -zN 250 -nX 129 -nY 129 -nZ 385 -dX 0.4 -dY 0.4 -dZ 0.4 -H -nosurf -pdb Molecule.pdb -inD RDCs.tbl -wv 0.12 -rM 8 -pka charges.pka -outD output.out, where PBLG.dx is the potential file of PBLG, Molecule.pdb is the PDB file of the studied molecule, RDCs.tbl is the list of experimental RDCs, charges.pka is the list of atomic charges obtained from AtomicChargeCalculator II [39], and output.out is the final output file. Atomic charge calculations were performed using the electronegativity equalisation method (EEM) [40] based on the atoms in molecules (AIM) calculation scheme at the B3LYP/6-311G level of theory [41].

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jmr.2022.107149.

References

[1] D. Chapman, A. Morrison, J. Biol. Chem. 241 (1966) 5044–5052.
[2] a) H. Kanoh, K. Yamada, F. Sakane, Trends Biochem. Sci. 15 (1990) 47–50; b) S.J. Ullrich, U.A. Hellmich, S. Ullrich, C. Glaubitz, Nat. Chem. Biol. 7 (2011) 263–270.
[3] N. Tsuji, J.L. Kennemur, T. Buyck, S. Lee, S. Prevost, P.S.J. Kaib, D. Bykov, C. Fares, B. List, Science 359 (2018) 1501–1505.
[4] M. Dutartre, J. Bayardon, S. Juge, Chem. Soc. Rev. 45 (2016) 5771–5794.
[5] G. Seydlová, R. Pohl, E. Zborníková, M. Ehn, O. Šimák, N. Panova, M. Kolář, K. Bogdanová, R. Večeřová, R. Bšer, H. Šanderová, D. Vítovská, P. Sudzinová, J. Pospíšil, O. Benada, T. Křížek, D. Sedlák, P. Bartůněk, L. Krásný, D. Rejman, J. Med. Chem. 60 (2017) 6098–6118.
[6] a) C. McGuigan, R.N. Pathirana, N. Mahmood, K.G. Devine, A.J. Hay, Antiviral Res. 17 (1992) 311–321; b) C. McGuigan, D. Cahard, H.M. Sheeka, E. De Clercq, J. Balzarini, J. Med. Chem. 39 (1996) 1748–1753; c) A.S. Alanazi, E. James, Y. Mehellou, ACS. Med. Chem. Lett. 10 (2019) 2–5.
[7] a) E. Procházková, J. Filo, M. Cigáň, O. Baszczyński, Eur. J. Org. Chem. 2020 (2020) 897–906; b) E. Procházková, P. Šimon, M. Straka, J. Filo, M. Májek, M. Cigáň, O. Baszczyński, Chem. Commun. 57 (2021) 211–214.
[8] P. Šimon, M. Tichotová, M.G. Gallardo, E. Procházková, O. Baszczyński, Chem. Eur. J. (2021).
[9] M. Ďud, M. Tichotová, E. Procházková, O. Baszczyński, Molecules 26 (2021) 5160.
[10] a) E. Procházková, R. Navrátil, Z. Janeba, J. Roithová, O. Baszczyński, Org. Biomol. Chem. 17 (2019) 315–320; b) A. Naito, T. Nagao, K. Norisada, T. Mizuno, S. Tuzi, H. Saito, Biophys. J. 78 (2000) 2405–2417.
[11] a) H. Hřebabecký, E. Procházková, M. Šála, P. Plačková, E. Tloušťová, O. Barauskas, Y.J. Lee, Y. Tian, R. Mackman, R. Nencka, Org. Biomol. Chem. 13 (2015) 9300–9313; b) E. Procházková, H. Hřebabecký, M. Dejmek, M. Šála, M. Šmídková, E. Tloušťová, E. Zborníková, L. Eyer, D. Růžek, R. Nencka, Bioorg. Med. Chem. Lett. 30 (2020) 1–4.
[12] a) L.D. Quin, M.J. Gallagher, G.T. Cunkle, D.B. Chesnut, J. Am. Chem. Soc. 102 (1980) 3136–3143; b) L. Benda, Z. Sochorová Vokáčová, M. Straka, V. Sychrovský, J. Phys. Chem. B 116 (2012) 3823–3833; c) I.B. Krivdin, Prog. Nucl. Magn. Reson. Spectrosc. 105 (2018) 54–99; d) G. Bifulco, R. Riccio, G.E. Martin, A.V. Buevich, R.T. Williamson, Org. Lett. 15 (2013) 654–657; e) R.T. Williamson, A.V. Buevich, G.E. Martin, Org. Lett. 14 (2012) 5098–5101; f) J.K. Harper, G.M. Pope, Annual Reports on NMR spectroscopy 98 (2019) 193–238.
[13] W. Kutzelnigg, U. Fleischer, M. Schindler, The IGLOMethod: Ab Initio Calculation and Interpretation of NMR Chemical Shifts and Magnetic Susceptibilities, NMR – Basis Principles and Progress, vol. 213, Springer, Berlin/Heidelberg, 1991.
[14] K. Wolinski, J.F. Hinton, P. Pulay, J. Am. Chem. Soc. 112 (1990) 8251–8260.
[15] J. Fukal, O. Páv, M. Buděšínský, I. Rosenberg, J. Šebera, V. Sychrovský, Phys. Chem. Chem. Phys. 21 (2019) 9924–9934.
[16] Z.J. He, Y.M. Wang, C.C. Tang, Phosphorus, Sulfur, and Silicon and the Related Elements 127 (1997) 59–66.
[17] a) A.L. Sklar, Rev. Mod. Phys. 14 (1942) 232; b) H.E. Smith, Chem. Rev. 98 (1998) 1709–1740.
[18] J.E. Kilpatrick, K.S. Pitzer, R. Spitzer, J. Am. Chem. Soc. 69 (1947) 2483–2488.
[19] C. Altona, H.J. Geise, C. Romers, Tetrahedron 24 (1968) 13–32.
[20] C. Altona, M. Sundaralingam, J. Am. Chem. Soc. 94 (1972) 8205–8212.
[21] J. Fukal, O. Páv, M. Buděšínský, J. Šebera, V. Sychrovský, Phys. Chem. Chem. Phys. 19 (2017) 31830–31841.
[22] A. Navarro-Vazquez, Magn. Reson. Chem. 50 (Suppl. 1) (2012) S73–S79.
[23] a) V. Schmidts, Ph.D. Thesis thesis, Technische Universität Darmstadt (Darmstadt), 2013; b) R. Berger, C. Fischer, M. Klessinger, J. Phys. Chem. A 102 (1998) 7157–7167.
[24] a) A. Ibanez de Opakua, F. Klama, I.E. Ndukwe, G.E. Martin, R.T. Williamson, M. Zweckstetter, Angew. Chem. Int. Ed. 59 (2020) 6172–6176; b) M. Zweckstetter, Nature Protocols 3 (2008) 679–690.
[25] R.S. Cahn, J. Chem. Educ. 41 (1964) 116–125.
[26] a) P.J. Stephens, F.J. Devlin, J.J. Pan, Chirality 20 (2008) 643–663; b) P.J. Stephens, D.M. McCann, F.J. Devlin, A.B. Smith 3rd, J. Nat. Prod. 69 (2006) 1055–1064.
[27] a) P. Tzvetkova, U. Sternberg, T. Gloge, A. Navarro-Vazquez, B. Luy, Chem. Sci. 10 (2019) 8774–8791; b) M.E. Di Pietro, U. Sternberg, B. Luy, J. Phys. Chem. B 123 (2019) 8480–8491.
[28] S. Parsons, H.D. Flack, T. Wagner, Acta Crystallogr. B 69 (2013) 249–259.
[29] a) G.M. Sheldrick, Acta Crystallogr. A 71 (2015) 3–8; b) G.M. Sheldrick, Acta Crystallogr. Section A 64 (2008) 112–122.
[30] G.M. Sheldrick, Acta Crystallogr. C 71 (2015) 3–8.
[31] Schrödinger release 2021-1 ed., Schrödinger, LLC, New York, 2021.
[32] C. Lu, C. Wu, D. Ghoreishi, W. Chen, L. Wang, W. Damm, G.A. Ross, M.K. Dahlgren, E. Russell, C.D. Von Bargen, R. Abel, R.A. Friesner, E.D. Harder, J. Chem. Theory Computation 17 (2021) 4291–4300.
[33] 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, Williams, 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, D.J. Fox, Gaussian, Inc., Wallingford, CT, 2016.
[34] a) A.D. Becke, J. Chem. Phys. 98 (1993) 5648-5652; b) C. Lee, W. Yang, R.G. Parr, Phys. Rev. B Condens Matter. 37 (1988) 785-789; c) S. H. Vosko, L. Wilk, M. Nusair, Can. J. Phys. 58 (1980) 1200-1211.
[35] V. Barone, M. Cossi, J. Phys. Chem. A 102 (1998) 1995–2001.
[36] S. Grimme, J. Antony, S. Ehrlich, H. Krieg, J. Chem. Phys. 132 (2010) 154104.
[37] M. Cossi, N. Rega, G. Scalmani, V. Barone, J. Comput. Chem. 24 (2003) 669–681.
[38] A. Ibáñez de Opakua, M. Zweckstetter, Magn. Reson. 2 (2021) 105–116.
[39] T. Raček, O. Schindler, D. Toušek, V. Horský, K. Berka, J. Koča, R. Svobodová, Nucleic Acids Res. 48 (2020) W591–W596.
[40] W.J. Mortier, S.K. Ghosh, S. Shankar, J. Am. Chem. Soc. 108 (1986) 4315–4320.
[41] S. Geidl, T. Bouchal, T. Raček, R. Svobodová Vareková, V. Hejret, A. Křenek, R. Abagyan, J. Koča, J. Cheminform. 7 (2015) 59.

IV.

M. Christou Tichotová, L. Tučková, H. Kocek, A. Růžička, M. Straka and E. Procházková:

Exploring the Impact of Alignment Media on RDC Analysis of Phosphorus-Containing Compounds: A Molecular Docking Approach

Check for updates

Exploring the impact of alignment media on RDC analysis of phosphorus-containing compounds: a molecular docking approach†

Markéta Christou Tichotová, [ab] Lucie Tučková, [a] Hugo Kocek, [a] Aleš Růžička, [c] Michal Straka [a] and Eliška Procházková [*a]

Residual dipolar couplings (RDCs) are employed in NMR analysis when conventional methods, such as J-couplings and nuclear Overhauser effects (NOEs) fail. Low-energy (optimized) conformers are often used as input structures in RDC analysis programs. However, these low-energy structures do not necessarily resemble conformations found in anisotropic environments due to interactions with the alignment medium, especially if the analyte molecules are flexible. Considering interactions with alignment media in RDC analysis, we developed and evaluated a molecular docking-based approach to generate more accurate conformer ensembles for compounds in the presence of the poly-γ-benzyl-ʟ-glutamate alignment medium. We designed chiral phosphorus-containing compounds that enabled us to utilize ^{31}P NMR parameters for the stereochemical analysis. Using P3D/PALES software to evaluate diastereomer discrimination, we found that our conformer ensembles outperform moderately the standard, low-energy conformers in RDC analysis. To further improve our results, we (i) averaged the experimental values of the molecular docking-based conformers by applying the Boltzmann distribution and (ii) optimized the structures through normal mode relaxation, thereby enhancing the Pearson correlation factor R and even diastereomer discrimination in some cases. Nevertheless, we presume that significant differences between J-couplings in isotropic and in anisotropic environments may preclude RDC measurements for flexible molecules. Therefore, generating conformer ensembles based on molecular docking enhances RDC analysis for mildly flexible systems while flexible molecules may require applying more advanced approaches, in particular approaches including dynamical effects.

Introduction

Conformation and configuration are key features of molecules, affecting their physical, chemical, and biological properties. Relative configuration can be determined by X-ray diffraction (XRD). However, XRD requires a single crystal, which is often difficult to obtain, particularly for biologically active compounds. Moreover, conformations in solid crystals do not necessarily match conformations in solution. In solution, molecular structures are typically analyzed by nuclear magnetic resonance (NMR) spectroscopy. Standard stereochemical analysis based on NMR

relies on vicinal ^{1}H–^{1}H ^{3}J-couplings, which are correlated to dihedral angles in a molecule through the Karplus equation,[1] or nuclear Overhauser effect[2,3] (NOE). When both J-couplings and NOEs fail, molecular stereochemistry may be analyzed using residual dipolar couplings (RDCs).

Residual dipolar couplings can be detected when molecular tumbling is reduced in an anisotropic environment.[4,5] The anisotropic environment is induced in alignment media,[6] such as poly-β-phenethylaspartates,[7] stretched polyvinyl acetate gels,[8] or poly-γ-benzyl-ʟ-glutamate[9] (PBLG). For structural determination, experimental RDC values are correlated with theoretical RDCs, which are back-calculated using alignment tensors of optimized molecular structures (conformers) of isomers studied.

Several program packages based on different principles are available to back-calculate theoretical RDCs (MSpin,[10] RDC@hotFCHT,[11] MDOC,[12] REDCAT,[13] and DipoCoup[14]). However, these approaches usually disregard specific analyte interactions with the alignment medium. To take these interactions into consideration, we must know the 3D structure of both interacting partners. However, to the best of our knowledge, the exact molecular structure of some

[a] Institute of Organic Chemistry and Biochemistry, Czech Academy of Sciences, 160 00 Prague, Czech Republic. E-mail: prochazkova@uochb.cas.cz

[b] Department of Physical and Macromolecular Chemistry, Faculty of Science, Charles University, 116 28 Prague, Czech Republic

[c] Department of General and Inorganic Chemistry, Faculty of Chemical Technology, University of Pardubice, Pardubice 532 10, Czech Republic

† Electronic supplementary information (ESI) available. CCDC 2277412–2277415. For ESI and crystallographic data in CIF or other electronic format see DOI: https://doi.org/10.1039/d3cp04099b

2016 | Phys. Chem. Chem. Phys., 2024, 26, 2016–2024

This journal is © the Owner Societies 2024

alignment media remains beyond the reach of contemporary experimentation.

P3D/PALES[15–17] was the first method to consider the atomic structure of the alignment medium in determining RDCs of small molecules. P3D/PALES has been adapted to a model of PBLG, which has a well-defined helical structure in chloroform.[18–20] The simulations use molecular structures of low-energy conformers as inputs (henceforth referred to as the low-energy approach). However, utilizing this approach in our previous work did not lead to the desired diastereomer discrimination, especially for flexible molecules.[21] The potential energy surface (PES) of flexible molecules is usually shallow and may contain several minima (corresponding to conformers) within a few tenths of kcal mol^{-1} from the global minimum (GM). Importantly, PES may be strongly affected by the environment, *i.e.*, by the presence of the alignment medium. Thus, low-energy structures provided for P3D/PALES simulations may significantly differ from those present in the experiment.

So far, several methods have been proposed for studying non-rigid molecules using RDCs. The most common RDC analysis relies on the alignment tensor fit; this has been adapted either as a single-alignment-tensor[22,23] or multiple-alignment-tensor[24,25] fitting approach to an optimised conformation ensemble. However, the idea of a single-alignment-tensor fitting to an ensemble of conformers is not generally applicable. The latter approach requires a large amount of anisotropic data, which are not always available. Another approach, usually used for macromolecules, employs the so-called reweight models[26–29] (maximum entropy, maximum parsimony, maximum allowed probability, *etc.*). However, these methods are computationally demanding and fail for more flexible systems. To overcome problems connected to the generation of structures, molecular-dynamics based MDOC[12,30,31] or multi-alignment-media based TITANIA[32] are an option. However, the main drawback of these methods is, again, the limited availability of larger amounts of experimental data.

The most widely used low-energy approach may not result in a correct back-calculation of theoretical RDCs of non-rigid molecules because a change in analyte conformation affects the RDC values. Based on this assumption, we raised the following research questions: (i) does the alignment medium significantly affect the analyte conformation? (ii) How large can the energy penalty paid by the analyte for the transition to the conformation constrained by the alignment medium be? We propose that conformer ensembles generated *via* molecular docking may provide a more accurate representation of the reality of the RDC experiment, consequently aiding in answering the posed questions.

Molecular docking is primarily used in drug design[33–35] to predict the binding pose (pose = conformation, position, and orientation) of a docked substrate (usually a small molecule = ligand) within a protein binding site. Nevertheless, by treating an analyte as the ligand and an alignment medium as the binding site, molecular docking may generate a more accurate conformer ensemble than the standard low-energy approach. Accordingly, we can consider the low-energy approach to generate conformers in isotropic solution and the molecular docking procedure to generate conformers in the presence of the alignment medium in RDC experiments.

In this study, we developed and evaluated a molecular docking-based approach to generate more accurate conformer ensembles for compounds in the poly-γ-benzyl-ʟ-glutamate (PBLG)[9,36–40] alignment medium environment. The model compounds used in this work feature a chiral phosphorus atom. We had shown in our previous work[21] that ^{31}P NMR parameters provide structural insights that aid the stereochemical analysis of phosphorus-containing compounds. We presume that the conformer ensembles generated *via* docking may provide a better approximation to the reality of RDC experiments than the conventional low-energy approach. Conformers generated by docking were used as inputs for RDC analysis. With the obtained back-calculated (theoretical) RDC datasets, we created Boltzmann distribution-weighted average RDC datasets which were then correlated with experimental RDC values. In addition, we relaxed the resulting geometries by constrained optimization in internal coordinates (henceforth referred to as *normal mode relaxation*). This approach preserves the overall molecular shape predicted by docking while relaxing the bond lengths and angles to obtain even more accurate input structures. The applicability of our molecular docking-based approach was evaluated by the P3D/PALES method. Finally, we also provide a comprehensive discussion of the underlying reasons behind the obtained results.

Results and discussion

Design and characterization of model compounds

We designed three classes of model compounds (Fig. 1) differing in molecular flexibility, as indicated by the nConf$_{20}$ parameter, which have been devised to quantify the flexibility of a molecule[41] (details are provided in the Experimental Section). As the most rigid compounds, we prepared phosphorylated derivatives of isopinocampheol (IPC; **1**, nConf$_{20}$ = 0), which is often used as a model compound in RDC studies.[9,42–44] We synthetized all four isomers of **1** to probe the robustness of the proposed docking-based approach. The absolute configuration of **1** was determined by X-ray diffraction (more details in ESI†). Subsequently, we applied the docking approach to mildly (**2**, nConf$_{20}$ = 2) and highly flexible compounds (**3**, nConf$_{20}$ = 8) that have been studied in our previous work.[21] Both **2** and **3** were characterized in our previous study[21] by means of NMR spectroscopy. Therein, also the absolute configuration of **2-SR** and **2-RR** was determined by X-ray diffraction. So far, we have been unable to prepare suitable crystals of **3** for X-ray diffraction analysis. This reluctance to crystallize has been previously associated with high molecular flexibility.[45]

Molecular docking of 1 (rigid structures)

Standard conformational analysis revealed only one low-energy conformer for each stereoisomer of **1**. Conversely, by molecular docking we found four unique conformers of **1-SR**, one of **1-SS**, three of **1-RS**, and three of **1-RR**. The DFT-calculated single

This journal is © the Owner Societies 2024

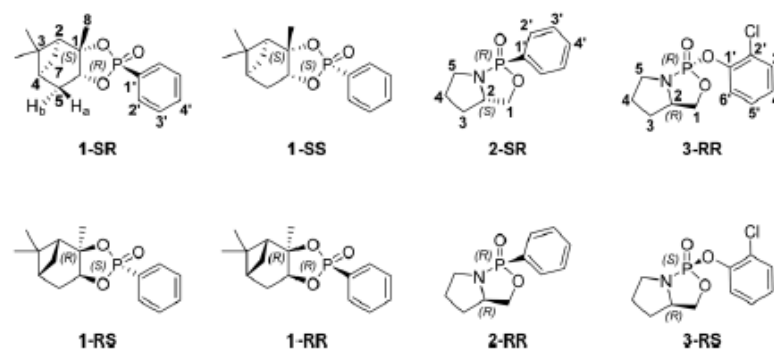*Phys. Chem. Chem. Phys.*, 2024, **26**, 2016–2024 | **2017**

Fig. 1 Chemical structures and atom numbering of the compounds studied. Absolute configuration of compounds **1** and **2** was determined by X-ray diffraction. Stereochemistry is denoted only for the relevant chiral centers (*i.e.*, phosphorus and carbon C1 for **1** or C2 for **2** and **3**).

point energies of the **1-SR** and **1-SS** conformers generated by docking were located approximately 6 kcal mol^{-1} above the individual global minima (GMs), whereas those of **1-RS** and **1-RR** were located more than 13 kcal mol^{-1} above GMs (Table S2 in ESI†).

The different energy distributions of unique conformations of individual isomers may be explained by differences in their diastereotopic interactions with the alignment medium. In particular, the energy penalty seems to be associated with the configuration on carbon C1 (Fig. 1). As shown in Fig. 2, the detailed microscopic view of poses obtained from molecular docking reveals that the (*SX*)-isomers (*X* = *R* or *S*; Fig. 2(a) and (b)) fit into the groove that winds around the PBLG polymer chain. By contrast, the (*RX*)-isomers (Fig. 2(c) and (d)) do not fit into this groove due to the shape of the bicyclic part of the molecule. The π–π stacking between the phenyl groups of the polymer and the analyte was observed

throughout majority of the poses. The distance between the stacked rings was ∼4 Å. Table S2 (ESI†) lists the DFT single-point energies of all unique conformers generated by docking and calculated relatively to the individual GMs ($\Delta E = E(\text{conf.}) - E(\text{GM})$, where E is the electronic energy of the given system).

To address the conformational tumbling of the molecule in the environment of the alignment medium, we calculated the Boltzmann distribution-weighted average of the RDCs of conformers obtained from docking. These RDCs (Unrelaxed. Boltz.) were correlated with all the experimental datasets of each isomer, as shown in Fig. 3 (the RDC values of each conformer are provided in Fig. S8–S23 in ESI†). The Pearson correlation coefficients (R) of the GM theoretical RDCs are also shown in Fig. 3.

For **1-SR** and **1-RS**, only the correct RDC datasets yields high correlation coefficients ($R > 0.8$); for **1-SS** and **1-RR**, both the
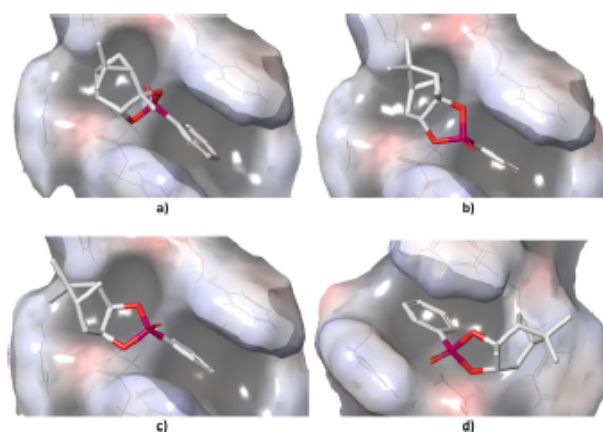

Fig. 2 Docked poses of (a) **1-SR**, (b) **1-SS**, (c) **1-RS**, and (d) **1-RR** in the helical structure of PBLG obtained from molecular docking. Carbon atoms are in white, oxygen in red, phosphorus in purple; hydrogens were omitted for clarity.
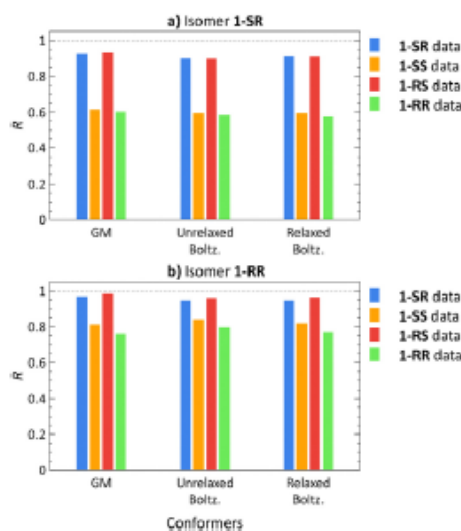
2018 | *Phys. Chem. Chem. Phys.*, 2024, **26**, 2016–2024

This journal is © the Owner Societies 2024

Fig. 3 Pearson correlation coefficients (R) of RDCs calculated using global minimum (GM) conformers, Boltzmann distribution-weighted theoretical RDCs of conformers obtained from molecular docking (Unrelaxed Boltz.), and the same conformers after normal mode relaxation (Relaxed Boltz.) of isomers 1-*SR* (a) and 1-*RR* (b) correlated with 1-*SR* (blue), 1-*SS* (orange), 1-*RS* (red), and 1-*RR* (green) RDC experimental datasets.

correct and the incorrect datasets lead to high correlation coefficients (Fig. 3). Nevertheless, the correct 1-*SS* and 1-*RR* datasets show a stronger correlation with the (correct) 1-*SS* and 1-*RR* structures, than with 1-*SR* and 1-*RS*. Therefore, diastereomer discrimination is not unequivocal although we are able to unilaterally determine the configuration of the enantiomer pairs of the compounds under study.

To improve the diastereomer discrimination, we relaxed the structures generated by docking using the normal mode relaxation method developed by Bouř *et al*[46,47] (computational details are provided in the Experimental Section). We assumed that density functional theory (DFT)-aided relaxation of the conformers obtained from docking would lead to more accurate conformer ensembles. The relaxed conformers were found approximately 1 kcal mol^{-1} for 1-*SR* and 1-*SS* and 2–4 kcal mol^{-1} for 1-*RS* and 1-*RR* above the corresponding GM (Table S3, ESI†). Normal mode relaxation slightly improved the diastereomer distinction, albeit with the same trends as those observed in unrelaxed conformers and GM structures. Thus, implementing molecular docking does not appear to be beneficial for RDC-aided structural elucidation of structures as rigid as isomers of 1 (nConf$_{20}$ = 0).

Molecular docking of 2 (mildly flexible molecules)

The molecules 2-*SR* and 2-*RR* are more flexible than the isomers of 1, as shown by the number of conformations (see below) as well as by nConf$_{20}$ parameter of 2, but still partly rigid

due to their two interconnected five-membered rings. Using the low-energy approach, we have previously identified three non-redundant low-energy conformers for each diastereomer of isolated system 2.[21] In turn, the docking procedure, which mimics the anisotropic environment, found tens of docking poses for each diastereomer of 2. Again, π–π stacking was observed for the majority of the discovered poses. However, many of the resulting conformers were nearly identical; hence, we had to apply a method for eliminating redundant conformers.

In our previous study, we have employed an approach based on dihedral angle analysis,[21] ultimately identifying three unique conformers for both isomers. In this work, we used a more complex machine learning (ML)-based algorithm for eliminating redundant conformers (henceforth referred to as ML elimination, see Experimental Section). As a result, we found 22 and 21 unique conformers for 2-*SR* and 2-*RR*, respectively. According to the DFT single-point energies, all of these conformers are located 4–10 kcal mol^{-1} above the corresponding GMs (Table S4, ESI†). The structures generated by docking were then relaxed by normal mode relaxation. The relaxed conformers lie 1–6 kcal mol^{-1} above the corresponding GMs (Table S5, ESI†).

Employing the Boltzmann distribution-weighted theoretical RDCs, we calculated the R values against the sets of experimental RDCs. The results shown in Fig. 4 demonstrate that the overall trends match those observed in our previous study[21] (also included in Fig. 4 for a direct comparison: conformers A, B, C).

The 2-*RR* RDC dataset provides the best diastereomer discrimination (Tables S57 and S59, ESI†), whereas the 2-*SR* dataset yields a good fit ($R > 0.8$) for both 2-*SR* and 2-*RR* structures (Tables S56 and S58, ESI†). However, the 2-*SR* structures show a higher overall correlation for conformers from molecular docking as compared to the low-energy approach. Furthermore, the additional normal mode relaxation increases the R values in all cases, except for the 2-*SR* RDC dataset with 2-*RR* conformers, where the R is slightly lower ($R = 0.988$ for the unrelaxed *versus* 0.985 for relaxed structures).

For the 2-*SR* isomer, the GM structure yields $R = 0.785$ using the 2-*SR* experimental RDCs,[21] Fig. 4. Employing conformers generated by molecular docking increases the 2-*SR* dataset R value to 0.815. Subsequent normal mode relaxation further improves the correlation ($R = 0.871$), albeit for both the correct and the incorrect datasets. Using conformers of 2-*RR*, diastereomer discrimination was not possible, neither using the low-energy nor the docking approach. Although the 2-*SR* dataset produces $R > 0.8$ for both 2-*RR* and 2-*SR* molecular structures, the 2-*RR* RDC dataset unambiguously yields a weak correlation for 2-*SR* ($R = 0.578$ after normal mode relaxation) and a strong correlation for 2-*RR* structures ($R = 0.964$ after normal mode relaxation). Thus, we have been able to achieve only unilateral diastereomer discrimination.

Notwithstanding the unilateral diastereomer discrimination, the structures can be assigned to datasets at a specific level of probability. Overall, the trends of 1 and 2 are rather similar. For visual comparison, the lowest energy conformers obtained in the previous study (A,[21] red), and the lowest energy
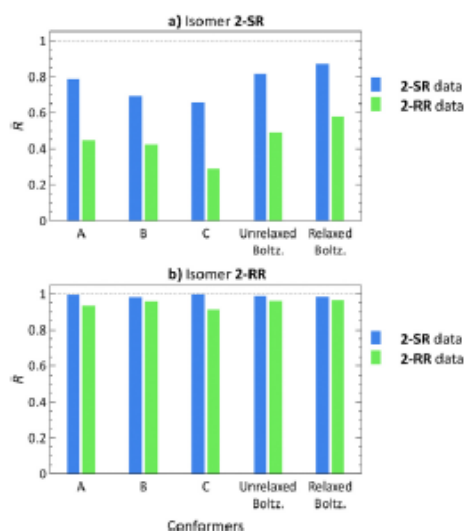
This journal is © the Owner Societies 2024

Phys. Chem. Chem. Phys., 2024, **26**, 2016–2024 | **2019**

Fig. 4 Pearson correlation coefficients (*R*) of RDCs calculated using low-energy conformers[21] (A)–(C), Boltzmann distribution-weighted theoretical RDCs of conformers obtained from molecular docking (Unrelaxed Boltz.), and the same conformers after relaxation by normal mode relaxation (Relaxed Boltz.) of structures **2-SR** (a) and **2-RR** (b) correlated with the **2-SR** (blue) and **2-RR** (green) RDC experimental datasets.

conformers generated by molecular docking with (green) and without (blue) subsequent normal mode relaxation are shown in Fig. 5.

Molecular docking of 3 (highly flexible molecules)

3-RS and **3-RR** are the most flexible molecules in this study. In our previous work,[21] conformational analysis of 3 with subsequent elimination of redundant structures produced 9 and 17 unique low-energy conformers for **3-RS** and **3-RR**, respectively. The docking approach with ML elimination of redundant conformers identified 23 unique conformers for **3-RS** and 34 conformers for **3-RR** isomer. The DFT-calculated single-point energies revealed that these conformers are located 6–15 kcal mol^{-1} above the

individual GMs, and 2–11 kcal mol^{-1} above the GMs after the subsequent normal mode optimization (details in Tables S6 and S7, ESI†). Unfortunately, we could not assign the experimental RDC datasets to the two isomers of 3, as we were unable to prepare crystals of 3 suitable for X-ray diffraction analysis (presumably because of the high flexibility of these molecules as indicated above). For this reason, the two measured RDC datasets are denoted as **3-A** and **3-B** in the following.

The **3-A** RDC dataset yields higher correlation coefficients with both **3-RS** and **3-RR** structures than the **3-B** dataset, Fig. 6. The best correlation coefficients ($R > 0.8$) were obtained for the normal mode-relaxed **3-RS** conformers with the **3-A** dataset. However, the **3-B** dataset fits on **3-RS** structures as well, although the absolute *R* values are smaller than for the **3-A** dataset ($R = 0.802$ and 0.590 for **3-A** and **3-B** RDC datasets, respectively). Thus, whether any degree of diastereomer discrimination can be inferred is at best questionable.

The trend, in which one dataset yields good correlation for both diastereomers was observed for all of the studied compounds, albeit less markedly for 3. So, does this trend have a physical meaning? We speculate that a possible explanation is that the isomers with an 'all-fitting' RDC dataset (**1-SR**, **1-RS**, **2-SR**, **3-A**) are moving more freely in the alignment medium environment. Thus, the induced RDCs are averaged over a higher number of geometries, accounting the goodness-of-fit of the experimental RDCs and the calculated RDCs of various conformers and isomers. Conversely, the isomers with the 'non-fitting' RDC dataset (**1-SS**, **1-RR**, **2-RR**, **3-B**) may be involved in isomer-specific interactions with the alignment medium, which cannot be adequately modelled by molecular docking. For 3, another possible explanation may be the higher variability of the RDC values of the **3-A** dataset (Tables S32–S39, ESI†). Consequently, the **3-A** dataset has higher *R* values than the **3-B** dataset.

Fig. 6 shows yet another trend: using normal mode relaxation significantly improves the correlation of the worst-fitting RDC dataset **3-B** with **3-RS** structures and worsens the correlation with **3-RR** isomer structures. A similar trend is observed in Fig. 4, where the **2-SR** RDC dataset shows a slightly lower *R* with the incorrect **2-RR** structures after normal mode relaxation, suggesting an incorrect structure/dataset match. However, without the XRD structure, we cannot ascertain whether this trend can help us to identify which dataset belongs to which isomer with a high level of certainty.
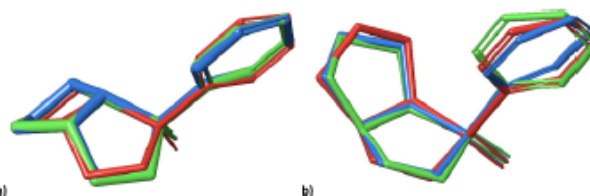

Fig. 5 Geometries of the lowest energy conformers of **2-SR** (a) and **2-RR** (b) from the previous study[21] (red), and the lowest energy conformers obtained from molecular docking into PLBG before (blue) and after normal mode relaxation (green).

2020 | *Phys. Chem. Chem. Phys.*, 2024, **26**, 2016–2024

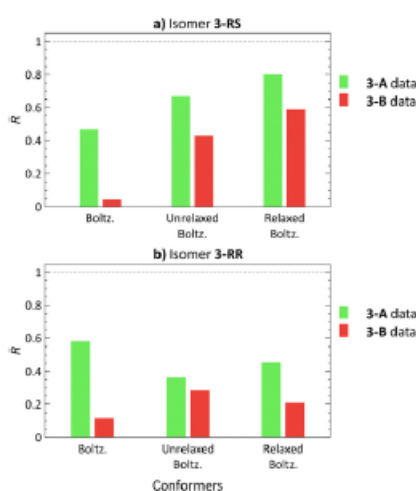This journal is © the Owner Societies 2024

Fig. 6 Pearson correlation coefficients (R) of Boltzmann distribution-weighted average RDCs calculated using theoretical RDCs of low-energy conformers[21] (Boltz.), conformers obtained from molecular docking (unrelaxed Boltz.) and these conformers after normal mode relaxation (relaxed Boltz.) of 3-RS (a) and 3-RR (b) correlated with the 3-A (green) and 3-B (red) RDC experimental datasets.

Furthermore, for highly flexible molecules such as 3, the conventional approach for the experimental determination of RDCs may fail.[12,24,32,48,49] Typically, the J-couplings extracted from experiment in the isotropic environment (i.e., isotropic J-couplings) are subtracted from the total T-couplings measured in the anisotropic environment to yield RDCs (eqn (1) in Experimental Section). However, the isotropic J-couplings may significantly differ from the real anisotropic J-couplings in the alignment medium, as indicated by the different conformer populations in isotropic and in anisotropic environments.

The differences between J-couplings calculated employing different conformer ensembles are clear and support our hypothesis: as shown in Table 1, we calculated J-couplings in the isotropic (presented by the low-energy conformer ensemble) and in the anisotropic environment (presented by the docking-obtained conformer ensemble). As an example, the calculated isotropic and anisotropic $J_{C1'-P}$-couplings of the 3-RR isomer are 3.8 and 1.0 Hz, respectively, resulting in a difference of 2.8 Hz between the isotropic and anisotropic value. For RDC calculated according to eqn (1) in Experimental Section, this difference would lead to an error of 2000% of the experimental RDC value determined using the isotropic J-coupling. More examples to show the difference between J-couplings calculated using different conformer ensembles can be found in Table 1. A more detailed study on this subject is currently underway.

Table 1 Comparison between experimental J-couplings of 3-A and 3-B measured in chloroform[21] (exp. J) and Boltzmann distribution-averaged J-couplings of 3-RS and 3-RR calculated at the B3LYP-D3/IGLO-III DFT level with PCM = chloroform of low-energy conformers (calc. J low-energy ensemble),[21] and relaxed conformers obtained from molecular docking (calc. J relaxed docking ensemble). Only calculated J-couplings of isomers with >2% relative abundance were used for the averaging

Interaction	Exp. J (Hz)[21]		Calc. J low-energy ensemble (Hz)[21]		Calc. J relaxed docking ensemble (Hz)	
	3-A	3-B	3-RS	3-RR	3-RS	3-RR
C1–P	4.8	4.6	6.8	5.6	6.9	5.1
C2–P	13.0	12.9	12.2	13.1	10.7	12.0
C3–P	4.1	3.6	3.1	3.8	0.4	1.0
C4–P	3.4	3.1	2.0	2.7	−0.7	0.8
C5–P	3.2	1.4	0.4	2.4	−0.5	1.4
C1'–P	6.1	8.9	−10.8	−10.7	−11.7	−7.7
C2'–P	7.6	8.9	8.7	5.9	10.9	7.0
C3'–P	2.9	2.6	−0.8	−2.3	0.0	−1.2
C4'–P	1.8	1.4	1.2	2.6	0.5	1.7
C5'–P	1.7	1.3	−1.5	−2.6	−0.8	−1.9
C6'–P	1.3	0.8	2.1	3.2	2.9	2.2
C2–H2	151.4	150.3	147.3	141.4	145.5	134.3
C3'–H3'	163.6	163.9	162.9	156.7	162.2	150.3
C4'–H4'	162.6	162.5	160.6	154.4	160.0	148.4
C5'–H5'	165.4	163.0	159.6	153.9	159.4	147.6
C6'–H6'	165.6	165.3	164.7	156.8	162.3	149.9

Experimental section

Computational methodology

Conformational sampling and molecular docking. As an input for molecular docking, structures generated via initial conformational sampling were used. The conformational sampling in MacroModel v13.1 (Schrödinger 2022-1 suite)[50] was performed with the following set of parameters: OPLS4 force field,[51] chloroform solvent, mixed torsional/low-mode sampling method with a maximum of 1000 steps, 40 kJ mol^{-1} energy window for saving structures, and 0.75 Å maximum atom deviation cut-off. Molecular docking was performed using Glide (as implemented in Schrödinger 2022-1 suite)[52–54] in flexible XP mode using grids of various sizes (10 Å, 12 Å, 15 Å, 18 Å, and 20 Å) to sample all possible positions of the analyte molecules. Aromatic hydrogen atoms were included as donors (0.0 partial charge cutoff), and chlorine was included as both the acceptor and the donor in separate searches with each grid for 3. The model of PBLG used in docking was provided to us by Alain Ibáñez de Opakua.[15]

Density functional theory (DFT) calculations. DFT calculations were performed using the Gaussian 16 program package.[55] Single-point energies of the conformers generated by molecular docking were calculated using the B3LYP functional[56–59] with empirical dispersion correction D3 by Grimme[60] and the 6-31+G(d,p)[61] basis set. DFT calculations were performed for single molecule in vacuo at 0 K. J-Couplings were calculated using the B3LYP functional and a basis set optimized for ^{31}P parameters (IGLO-II)[62–64] within the polarizable continuum model (PCM)[65] using chloroform as solvent.

Machine learning-aided elimination of redundant structures. Unique conformers were identified using the machine learning (ML)-based mean-shift algorithm[66] as implemented in

This journal is © the Owner Societies 2024

Phys. Chem. Chem. Phys., 2024, 26, 2016–2024 | 2021

the Scikit-learn Python package (scikit-learn.org). This approach was previously adapted for identification of non-redundant conformers in a series of dipeptides.[67] During the ML-aided elimination procedure, the conformers were sorted into clusters based on selected dihedral angles. Each cluster contained all structurally similar conformers. The mean shift clustering algorithm specifies the cluster size using a bandwidth parameter. The value of this parameter determines whether each data point belongs to a given cluster, i.e., if a given set of conformers is structurally similar. This parameter is proportional to the size of a given cluster. The optimal values for each structure were decided as described in the ESI† (numbers are provided in Table S64). For each cluster, only the lowest energy conformer was selected as the unique non-redundant conformer.

Normal mode relaxation. Normal mode optimization has been used to refine molecular structures, namely bond lengths and angles, while keeping the overall shape of the molecule, i.e., by fixing the lowest vibrational modes.[46,47] In this procedure, modes up to 300 cm^{-1} were fixed, and the QGRAD program was used for normal mode optimization. The accompanying DFT computations were performed at the B3LYP-D3/DGDZVP[68] level using the Gaussian 16 software suite.[55] The final energies of the relaxed conformers were recalculated at the B3LYP-D3/6-31+G(d,p) level.

P3D/PALES. We performed molecular alignment simulations using P3D[15] as implemented in the PALES[17] software, as recommended.[16] The following command was used to run the simulation:

pales -elPales -3D -pot3D PBLG.dx -lcS 0.8 -maxPot 2 -z1

150 -zN 250 -nX 129 -nY 129 -nZ 385 -dX 0.4 -dY 0.4 -dZ

0.4 -H -nosurf -pdb Molecule.pdb -inD RDCs.tbl -wv

0.12 -rM 8 -pka charges.pka -outD output.out,

where PBLG.dx is the potential file of PBLG, Molecule.pdb is the PDB file of the studied molecule, RDCs.tbl is the list of experimental RDCs, charges.pka is the list of atomic charges obtained from AtomicChargeCalculator II,[69] and output.out is the final output file. Atomic charges were calculated using the electronegativity equalization method (EEM)[70] based on the atoms in molecules (AIM) calculation scheme at the B3LYP/6-311G level of theory.[71]

Calculation of nConf$_{20}$. We calculated the values of the nConf$_{20}$ flexibility descriptor using the code provided in the original article by Wicker and Cooper.[41] Given that a more detailed scale was necessary for the studied compounds, we set the atom root-mean-square (RMS) distance threshold for removing duplicate conformers to 0.5 Å. The number of generated conformers was set to 10 000 to mitigate the influence of stochasticity in the process of structure generation.

Experimental methodology

Sample preparation. The quantity of the alignment media was derived from the quantities of the solvent and analyte to acquire approximately 6.1–7.9 weight% of the alignment medium in the sample. This number was determined based on our

experience[21,46,72] and on the premeasurement screening. The calculated amount of the alignment medium was mixed with the respective amounts of the solvent and the analyte (detailed sample compositions are provided in ESI†), and the final mixture was left standing to dissolve overnight.

Due to the high viscosity of the solution, the sample was homogenized using a manual centrifuge. Sample homogeneity was then monitored by ^{2}H image experiments, and the alignment order was determined by measuring the quadrupolar splitting of the solvent in ^{2}H NMR spectra.

Extraction of experimental RDCs and their evaluation. Residual dipolar couplings were acquired using F1-coupled HSQC (one-bond ^{13}C–^{1}H couplings) or APT (^{13}C–^{31}P couplings) experiments. The F1-coupled HSQC sequence was J-scaled, and the separation of the components was 8-times enhanced with respect to the ^{13}C chemical shift evolution.[44] The one-bond ^{13}C–^{1}H and ^{13}C–^{31}P couplings were obtained using the following equation:

$$^nD_{C-Y} = \frac{^nT_{C-Y} - ^nJ_{C-Y}}{2}, \qquad (1)$$

where J stands for the scalar coupling (obtained from an isotropic experiment – measured in CDCl$_3$), T is the total coupling (collected from an anisotropic experiment measured in the solvent CDCl$_3$ and in the alignment medium PBLG), D is the residual dipolar coupling, index ^{n} indicates the number of bonds, and C–Y stands for ^{13}C–^{1}H or ^{13}C–^{31}P coupling. The values corresponding to the rotating t-butyl groups were converted into one-bond C–C $^1D_{CC}$ – the RDC between methyl and tertiary carbon nuclei.[73]

The RDC data were evaluated using Pearson's correlation coefficient R and quality factor Q.

Pearson's correlation coefficient R is defined as follows:

$$R = \frac{n\sum D_{exp}D_{calc} - \sum D_{exp}\sum D_{calc}}{\sqrt{n\sum D_{exp}^2 - \left(\sum D_{exp}\right)^2}\sqrt{n\sum D_{calc}^2 - \left(\sum D_{calc}\right)^2}}, \qquad (2)$$

where n is the number of experimental RDC values, D_{exp} are the experimental RDCs, and D_{calc} are the theoretical RDC values.

The quality factor is given by the following equation:

$$Q = \sqrt{\frac{\frac{1}{n}\sum\left(D_{exp}^2 - D_{calc}^2\right)}{\frac{1}{n}\sum D_{exp}^2}}, \qquad (3)$$

where D_{exp} are the experimental RDC values, D_{calc} are the calculated RDCs, and n is the number of experimental RDCs.

Conclusions

In this work, we investigated whether the involvement of alignment media in generation of conformer ensembles via molecular docking improves the performance of RDC-aided stereochemical analysis. We have found that it depends on the flexibility of the molecules under study. The docking approach does not markedly affect the conformation space of rigid molecules such as 1, neither it significantly improves the quality of RDC analysis. However, the different conformer ensembles do affect the results, indicating that the alignment medium influences the conformation space of

2022 | Phys. Chem. Chem. Phys., 2024, 26, 2016–2024

This journal is © the Owner Societies 2024

the analyte molecules. Generally, a non-negligible energetic penalty for higher-lying conformers may be compensated for by the analyte–medium interactions which may extend/alter the conformer ensembles in the anisotropic environment in comparison with the isotropic solution. This energy penalty is diastereotopic (isomer-dependent), this effect is most profound in the case of rigid molecules, such as isomers of **1** that are not as capable of adapting to the chiral environment of PBLG. For mildly flexible systems, such as **2**, the ensembles of conformers generated by molecular docking clearly enhance the accuracy of RDC-aided stereochemical analysis. Additional DFT optimization of conformers obtained from molecular docking by constrained normal mode relaxation further improves the accuracy of RDC analysis and brings in a clearer diastereomer discrimination. The docking approach does not improve the results of RDC analysis for highly flexible structures, such as **3**; we speculate that this is due to the possibly incorrect experimental determination of RDCs using J-couplings obtained from experiments in isotropic solutions. The total coupling T of an anisotropic conformer ensemble consists of RDCs and J-couplings, however, the difference between anisotropic an isotropic J-couplings may increase with molecular flexibility due to the fact that J-couplings are rather sensitive to changes in molecular geometries. This difference may preclude RDC determination for highly flexible molecules because – like RDCs – anisotropic J-couplings cannot be measured directly. To answer the questions laid in the Introduction, the consideration of alignment medium does have an effect on the conformation space of analyte and thus the quality of RDC analysis. Even more accurate conformer ensembles may be generated using advanced simulation methods (employing, *e.g.*, QM/MM or molecular dynamics). Furthermore, the simple model containing "static snapshots", *i.e.*, singular conformers of the studied systems, is perhaps insufficient for the correct description of the entire conformational space. Methods addressing fast molecular motion and dynamic effects may provide more precise results, however, at a larger computational cost. From experimental point of view, the effect of different alignment media on the RDC values should be compared because specific analyte–medium interactions likely induce different conformer ensembles, which are associated with different sets of RDCs.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

References

1 M. Karplus, *J. Am. Chem. Soc.*, 1963, **85**, 2870–2871.

2 R. A. Bell and J. K. Saunders, *Can. J. Chem.*, 1970, **48**, 1114–1122.

3 A. W. Overhauser, *Phys. Rev.*, 1953, **92**, 411–415.

4 F. Kramer, M. V. Deshmukh, H. Kessler and S. J. Glaser, *Concepts Magn. Reson., Part A*, 2004, **21A**, 10–21.

5 C. M. Thiele, *Eur. J. Org. Chem.*, 2008, 5673–5685.

6 G. Kummerlöwe and B. Luy, *TrAC, Trends Anal. Chem.*, 2009, **28**, 483–493.

7 M. Schwab, D. Herold and C. M. Thiele, *Chem. – Eur. J.*, 2017, **23**, 14576–14584.

8 J. C. Freudenberger, S. Knor, K. Kobzar, D. Heckmann, T. Paululat, H. Kessler and B. Luy, *Angew. Chem., Int. Ed.*, 2005, **44**, 423–426.

9 A. Marx and C. Thiele, *Chem. – Eur. J.*, 2009, **15**, 254–260.

10 A. Navarro-Vázquez, *Magn. Reson. Chem.*, 2012, **50**, S73–S79.

11 V. Schmidts, *PhD thesis*, Technische Universität Darmstadt, 2013.

12 M. Di Pietro, U. Sternberg and B. Luy, *J. Phys. Chem. B*, 2019, **123**, 8480–8491.

13 H. Valafar and J. H. Prestegard, *J. Magn. Reson.*, 2004, **167**, 228–241.

14 J. Meiler, W. Peti and C. Griesinger, *J. Biomol. NMR*, 2000, **17**, 283–294.

15 A. Ibáñez de Opakua, F. Klama, I. E. Ndukwe, G. E. Martin, R. T. Williamson and M. Zweckstetter, *Angew. Chem., Int. Ed.*, 2020, **59**, 6172–6176.

16 A. Ibáñez de Opakua and M. Zweckstetter, *Magn. Reson.*, 2021, **2**, 105–116.

17 M. Zweckstetter, *Nat. Protoc.*, 2008, **3**, 679–690.

18 J. R. Kim and T. Ree, *J. Polym. Sci., Polym. Chem. Ed.*, 1985, **23**, 215–221.

19 P. Doty, J. H. Bradbury and A. M. Holtzer, *J. Am. Chem. Soc.*, 1956, **78**, 947–954.

20 M. Sarfati, P. Lesot, D. Merlet and J. Courtieu, *Chem. Commun.*, 2000, 2069–2081.

21 M. Tichotová, A. Ešnerová, L. Tučková, L. Bednárová, I. Císařová, O. Baszczyński and E. Procházková, *J. Magn. Reson.*, 2022, **336**, 107149.

22 A. Schuetz, J. Junker, A. Leonov, O. F. Lange, T. F. Molinski and C. Griesinger, *J. Am. Chem. Soc.*, 2007, **129**, 15114–15115.

23 P. Trigo-Mouriño, R. Santamaría-Fernández, V. M. Sánchez-Pedregal and A. Navarro-Vázquez, *J. Org. Chem.*, 2010, **75**, 3101–3104.

24 C. M. Thiele, V. Schmidts, B. Böttcher, I. Louzao, R. Berger, A. Maliniak and B. Stevensson, *Angew. Chem., Int. Ed.*, 2009, **48**, 6708–6712.

25 C. M. Thiele, A. Maliniak and B. Stevensson, *J. Am. Chem. Soc.*, 2009, **131**, 12878–12879.

26 D. Selegato, C. Bracco, C. Giannelli, G. Parigi, C. Luchinat, L. Sgheri and E. Ravera, *ChemPhysChem*, 2021, **22**, 127–138.

27 E. Ravera, L. Sgheri, G. Parigi and C. Luchinat, *Phys. Chem. Chem. Phys.*, 2016, **18**, 5686–5701.

28 S. Bottaro, T. Bengtsen and K. Lindorff-Larsen, *Methods Mol. Biol.*, 2020, **2112**, 219–240.

29 I. Bertini, Y. K. Gupta, C. Luchinat, G. Parigi, M. Peana, L. Sgheri and J. Yuan, *J. Am. Chem. Soc.*, 2007, **129**, 12786–12794.

This journal is © the Owner Societies 2024

Phys. Chem. Chem. Phys., 2024, **26**, 2016–2024 | **2023**

30 U. Sternberg, P. Tzvetkova and C. Muhle-Goll, *Phys. Chem. Chem. Phys.*, 2020, **22**, 17375–17384.

31 P. Tzvetkova, U. Sternberg, T. Gloge, A. Vazquez and B. Luy, *Chem. Sci.*, 2019, **10**, 8774–8791.

32 F. A. Roth, V. Schmidts, J. Rettig and C. M. Thiele, *Phys. Chem. Chem. Phys.*, 2022, **24**, 281–286.

33 F. Stanzione, I. Giangreco and J. C. Cole, *Prog. Med. Chem.*, 2021, **60**, 273–343.

34 N. S. Pagadala, K. Syed and J. Tuszynski, *Biophys. Rev.*, 2017, **9**, 91–102.

35 K. A. Peele, C. Potla Durthi, T. Srihansa, S. Krupanidhi, V. S. Ayyagari, D. J. Babu, M. Indira, A. R. Reddy and T. C. Venkateswarulu, *Inform. Med. Unlocked*, 2020, **19**, 100345.

36 V. M. Marathias, G. J. Tawa, I. Goljer and A. C. Bach, *Chirality*, 2007, **19**, 741–750.

37 A. Marx, V. Schmidts and C. M. Thiele, *Magn. Reson. Chem.*, 2009, **47**, 734–740.

38 A. Marx, B. Böttcher and C. M. Thiele, *Chem. – Eur. J.*, 2010, **16**, 1656–1663.

39 L. Vrzal, M. Kratochvilova-Simanova, T. Landovsky, K. Polivkova, J. Budka, H. Dvorakova and P. Lhotak, *Org. Biomol. Chem.*, 2015, **13**, 9610–9618.

40 T. Landovský, M. Tichotová, L. Vrzal, J. Budka, V. Eigner, H. Dvořáková and P. Lhoták, *Tetrahedron*, 2018, **74**, 902–907.

41 J. G. P. Wicker and R. I. Cooper, *J. Chem. Inf. Model.*, 2016, **56**, 2347–2352.

42 S. Hansmann, T. Larem and C. M. Thiele, *Eur. J. Org. Chem.*, 2016, 1324–1329.

43 T. Montag and C. M. Thiele, *Chem. – Eur. J.*, 2013, **19**, 2271–2274.

44 S. Jeziorowski and C. Thiele, *Chem. – Eur. J.*, 2018, **24**, 15631–15637.

45 L. Yu, S. M. Reutzel-Edens and C. A. Mitchell, *Org. Process Res. Dev.*, 2000, **4**, 396–402.

46 P. Bouř and T. A. Keiderling, *J. Chem. Phys.*, 2002, **117**, 4126–4132.

47 P. Bouř, *Collect. Czech. Chem. Commun.*, 2005, **70**, 1315–1340.

48 A. Canales, J. Jiménez-Barbero and M. Martin-Pastor, *Magn. Reson. Chem.*, 2012, **50**, S80–S85.

49 M. E. Di Pietro, G. De Luca, G. Celebre, D. Merlet and C. Aroulanda, *Mol. Cryst. Liq. Cryst.*, 2015, **614**, 39–53.

50 *Schrödinger Release 2022-1 ed.*, Schrödinger, LLC, New York, NY, 2022.

51 C. Lu, C. Wu, D. Ghoreishi, W. Chen, L. Wang, W. Damm, G. A. Ross, M. K. Dahlgren, E. Russell, C. D. Von Bargen, R. Abel, R. A. Friesner and E. D. Harder, *J. Chem. Theory Comput.*, 2021, **17**, 4291–4300.

52 R. A. Friesner, R. B. Murphy, M. P. Repasky, L. L. Frye, J. R. Greenwood, T. A. Halgren, P. C. Sanschagrin and D. T. Mainz, *J. Med. Chem.*, 2006, **49**, 6177–6196.

53 T. A. Halgren, R. B. Murphy, R. A. Friesner, H. S. Beard, L. L. Frye, W. T. Pollard and J. L. Banks, *J. Med. Chem.*, 2004, **47**, 1750–1759.

54 R. A. Friesner, J. L. Banks, R. B. Murphy, T. A. Halgren, J. J. Klicic, D. T. Mainz, M. P. Repasky, E. H. Knoll, M. Shelley, J. K. Perry, D. E. Shaw, P. Francis and P. S. Shenkin, *J. Med. Chem.*, 2004, **47**, 1739–1749.

55 *Gaussian 16 Rev. C.01*, Wallingford, CT, 2016.

56 C. Lee, W. Yang and R. G. Parr, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1988, **37**, 785–789.

57 A. D. Becke, *J. Chem. Phys.*, 1993, **98**, 5648–5652.

58 S. H. Vosko, L. Wilk and M. Nusair, *Can. J. Phys.*, 1980, **58**, 1200–1211.

59 P. J. Stephens, F. J. Devlin, C. F. Chabalowski and M. J. Frisch, *J. Phys. Chem.*, 1994, **98**, 11623–11627.

60 S. Grimme, J. Antony, S. Ehrlich and H. Krieg, *J. Chem. Phys.*, 2010, **132**, 154104.

61 V. Barone and M. Cossi, *J. Phys. Chem. A*, 1998, **102**, 1995–2001.

62 L. Benda, Z. Sochorová Vokáčová, M. Straka and V. Sychrovský, *J. Phys. Chem. B*, 2012, **116**, 3823–3833.

63 J. Fukal, O. Páv, M. Buděšínský, I. Rosenberg, J. Šebera and V. Sychrovský, *Phys. Chem. Chem. Phys.*, 2019, **21**, 9924–9934.

64 J. Fukal, O. Páv, M. Buděšínský, J. Šebera and V. Sychrovský, *Phys. Chem. Chem. Phys.*, 2017, **19**, 31830–31841.

65 M. Cossi, N. Rega, G. Scalmani and V. Barone, *J. Comput. Chem.*, 2003, **24**, 669–681.

66 D. Comaniciu and P. Meer, *IEEE Trans. Pattern Anal. Mach. Intell.*, 2002, **24**, 603–619.

67 T. Kalvoda, M. Culka, L. Rulíšek and E. Andris, *J. Phys. Chem. B*, 2022, **126**, 5949–5958.

68 N. Godbout, D. Salahub, J. Andzelm and E. Wimmer, *Can. J. Chem.*, 2011, **70**, 560–571.

69 T. Raček, O. Schindler, D. Toušek, V. Horský, K. Berka, J. Koča and R. Svobodová, *Nucleic Acids Res.*, 2020, **48**, W591–W596.

70 W. J. Mortier, S. K. Ghosh and S. Shankar, *J. Am. Chem. Soc.*, 1986, **108**, 4315–4320.

71 S. Geidl, T. Bouchal, T. Raček, R. Svobodová Vařeková, V. Hejret, A. Křenek, R. Abagyan and J. Koča, *J. Cheminformatics*, 2015, **7**, 59.

72 M. Tichotová, T. Landovský, J. Lang, S. Jeziorowski, V. Schmidts, M. Kohout, M. Babor, P. Lhoták, C. M. Thiele and H. Dvořáková, *J. Org. Chem.*, 2023, DOI: 10.1021/acs.joc.2c02594.

73 R. S. Stoll, M. V. Peters, A. Kuhn, S. Heiles, R. Goddard, M. Bühl, C. M. Thiele and S. Hecht, *J. Am. Chem. Soc.*, 2009, **131**, 357–367.

V.

U. Sternberg, **M. Christou Tichotová**, L. Tučková, A. Ešnerová, J. Hanus, O. Baszczyňski and E. Procházková:

Extending Molecular Dynamics with Dipolar NMR Tensors as Constraints to Chiral Phosphorus Compounds

Extending Molecular Dynamics with Dipolar NMR Tensors as Constraints to Chiral Phosphorus Compounds

Ulrich Sternberg,[a,b] Markéta Christou Tichotová,[c,d] Lucie Tučková,[c] Aneta Ešnerová,[e] Jan Hanus,[c] Ondřej Baszczyński,[c,e] and Eliška Procházková[c,*]

Molecular Dynamics with Orientational Constraints (MDOC) simulations use NMR parameters as tensorial constraints in the stereochemical analysis of small molecules. ^{13}C-^{31}P Residual dipolar couplings-aided MDOC simulations of small phosphorus molecules determined the relative configurations of rigid molecules after including $^3J_{H-H}$-couplings as additional constraints. However, flexible molecules remain a problem.

Introduction

Compounds with the stereogenic centre on a phosphorus atom (e.g., Josiphos ligands[1]) are essential in coordination chemistry and stereoselective catalysis. However, P-chiral compounds are often difficult to synthesize stereoselectively, and classical synthesis yields a mixture of isomers which may be difficult to separate.[2,3] To determine phosphorus stereochemistry (configuration and conformation), several advanced NMR methods can be used, such as residual dipolar coupling (RDC) analysis.[4-8]

The RDC analysis enables to assign the relative configuration of small to medium sized molecules,[7-11] but some problems arise upon high molecular flexibility[12-15] or lack of ^1H–^{13}C RDC values. Previously,[16] we have studied model phosphorus-based compounds by ^{13}C–^{31}P RDC analysis, with the phosphorus atom incorporated into a cycle to suppress molecular flexibility. However, we were unable to unambiguously determine the relative configuration using the P-based RDC analysis of low-energy conformers. These low-energy structures probably do not reflect the actual structures present in the anisotropic environment of RDC experiments. To create a new population of conformers in the alignment medium needed for RDC analysis, we applied molecular docking.[17,18]

[a] COSMOS-Software, 07743 Jena, Germany
[b] Karlsruhe Institute of Technology (KIT), Postfach 3640, D-76021 Karlsruhe, Germany
[c] Institute of Organic Chemistry and Biochemistry, Czech Academy of Sciences, Prague 166 10, Czech Republic
[d] Department of Physical and Macromolecular Chemistry, Faculty of Science, Charles University, Prague 128 43, Czech Republic
[e] Department of Organic Chemistry, Faculty of Science, Charles University, Prague 128 43, Czech Republic
† Footnotes relating to the title and/or authors should appear here.
Supplementary Information available: [details of any supplementary information available should be included here]. See DOI: 10.1039/x0xx00000x
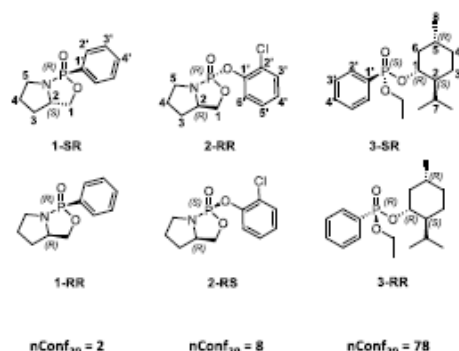
Molecular docking ultimately improved our ability to determine the relative configuration of rigid molecules, but not highly flexible molecules.[19] Nevertheless, this issue may be solved using a recently developed method: Molecular Dynamics with Orientational Constraints (MDOC).[20-22]

MDOC generates trajectories comprising a multitude of conformers using dipolar-coupling tensors (or other NMR interaction tensors) for each H–C or C–P coupling. To reach the time scale of NMR experiments in molecular dynamics simulations, MDOC applies tensorial constraints that rotate molecules and their mobile groups, thus making it possible to heat up the rotational degrees of freedom. Moreover, we can adjust equilibria between rotamers using scalar constraints, such as J-couplings and NOE distances.

MDOC has been recently applied to the structural elucidation and diastereomer discrimination of several small and flexible molecules, such as cellobiose[22] and oidiolactone B,[21] and of molecules with multiple rotamers, such as sagittamide A.[23] However, a recent MDOC study[24] has highlighted two major issues that must be considered in stereochemical analysis: (i) stiff molecules, e.g., strychnine, actually have several conformers in solution, and (ii) transient structures, mainly controlled by the entropy term $T\Delta S$ of free energy, also contribute to conformer distribution.

In this study, we applied MDOC to model phosphorus-containing compounds (Fig. 1) with various degrees of molecular flexibility, as indicated by the nConf$_{20}$ parameter. This parameter can be used for quantification of molecular flexibility[25] (calculation details are provided in the ESI). The descriptor is the count of conformers of a molecule with energies between the lowest energy conformer and the selected relative energy threshold and presents consequently the number of energetically accessible conformations.

1-SR   2-RR   3-SR

1-RR   2-RS   3-RR

$nConf_{20} = 2$ $nConf_{20} = 8$ $nConf_{20} = 78$

Fig. 1: Structures of the study compounds and their flexibility parameters.

Fig. 2: Statistical analysis of NMR data calculated from MDOC trajectories of 1.

In this work, we implemented ^{31}P NMR parameters (RDCs and J-couplings) in MDOC analysis, showing that these data may facilitate stereochemical analysis of phosphorus compounds.

Results and Discussion

MDOC analysis of 1

Diastereomers of **1** were prepared as described in our previous study and the absolute configuration was determined by X-ray diffraction.[16] In the present investigation, a 40 ns MDOC simulation was performed for both stereoisomers with tha data sets 1A and 1B (for the details parameters of the simulation see the ESI).

For both datasets, 1A and 1B, the quality factors n/$\chi 2$[26] (details in the section "Computational Methods") were low, reaching 0.1 and 0.2, respectively, due to the relatively small $^{13}C-^{31}P$ couplings and narrow error ranges. Because these quality factor values prevented us from assigning the relative configuration, we calculated the χ-probability[26] (Table S10) and assigned dataset 1A as **1-SR** by 90 %, in line with X-ray data.[16] However, the **1-RR** configuration was still unclear. As in our previous studies,[16, 19] we were unable to identify the relative configuration using only RDC data (Table S10). For this reason, we included J-couplings in MDOC simulations.

J-couplings are valuable parameters, especially for stereogenic centres, but we failed to experimentally assign the $^{3}J_{H-H}$-couplings of the diastereotopic hydrogens H1A (proS) and H1B (proR) to H2 (see Fig. 3). So, we performed preliminary MDOC simulations, but instead of using the $^{3}J_{H-H}$-couplings as additional constraints, we calculated them as the mean value of 2000 snapshots of the MDOC trajectory. With these computed values, we assigned $^{3}J_{H-H}$-couplings (Table S13).

In the final MDOC simulation, using $^{3}J_{H-H}$-couplings as additional scalar constraints, we reached 96 and 97% χ-probability of assigning **1-SR** and **1-RR** to datasets 1A and 1B, respectively (Fig. 2).
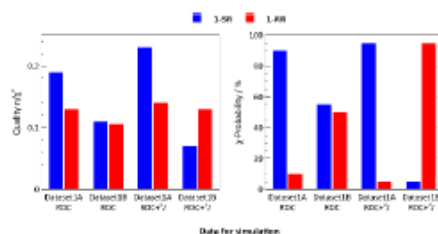
Encouraged by these results, we investigated conformational equilibria from the 2000 snapshots of the trajectories. From MDOC trajectories of the torsion angles of **1-SR**, we derived the populations of individual conformers. The conformations of the five-membered rings were investigated using the pseudorotational phase angle P, as introduced by Altona and Sundaralingami.[27] The maximum of the distribution was at $P = 208°$, and most (98.4 %) P values belonged to the South conformation. In this simulation, we identified another five-membered ring conformation separated from the major conformations by a minimum. However, this North conformation accounted for only 1.6% of occurrences (see Fig. 3). Inspecting the torsion of the (O2–**C1**–**C2**–N1) angle of the adjacent five-membered ring (see Fig.3) we observe a small -gauche contribution of about 6% additionally to the +gauche conformation, with a maximum P at 28° (Fig. 3). The two adjacent ring systems preform a coupled motion.

The dihedral distributions of **1-RR** were like those of **1-SR** (Fig. 3), but the maximum of the (O2–**C1**–**C2**–N1) distribution was at 20° (+gauche in 96 %). And while we also observed two conformers of the five-membered rings, the contribution of the second conformer was minimal (for details see Tab. 1). The maximum of the pseudo-rotational angle of the more populated conformation (98.2 %) was at $P = 168°$. As with **1-SR**, the phenyl rotation was not restricted.
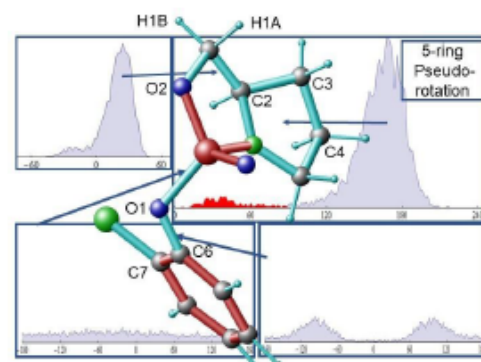

Fig. 3: Distribution of torsion angles of 1-SR as derived from a MDOC trajectory; the (N1–P1–C6–C7) torsion (upper panel) indicates that the rotation of the phenyl group is not restricted (for details, see Tab. 1). The conformers of the five-membered rings containing the phosphorous atom are displayed using the rotation about the C1-C2 bond (lower left panel) and the conformation of the adjacent 5-ring is analysed using the pseudo-rotation angle (right panel).
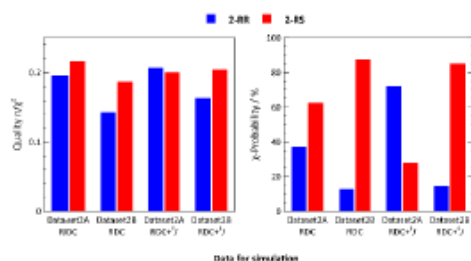
Fig. 4: Statistical analysis of NMR data calculated from MDOC trajectories of 2.
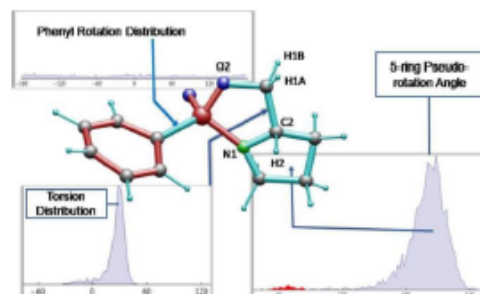

Fig. 5: Distribution of torsion angles of 2-RR calculated from the MDOC trajectory; the chlorophenyl residue is linked to the phosphorous over an oxygen bridge. The (N1–P1–O1–C6) torsion (lower left panel) indicates that the rotation is not restricted but the rotation about the (P1-O1-C6-C7) bond displays two minima (lower right panel. The conformers of the five-membered rings containing the phosphorous atom is displayed using the rotation about the C1-C2 bond (upper left panel) and the conformation of the adjacent 5-ring is analysed using the pseudo-rotation angle (right panel) (for details, see Tab. 1).

MDOC analysis of 2

Diastereomers of **2** were prepared as described in our previous study.[16] However, crystallization always caused ring opening leading to loss of the stereogenic centre on the phosphorus atom, therefore, absolute configuration was not determined. For this reason, we worked with both NMR datasets 2A and 2B. The MDOC analysis of **2** showed that the quality n/χ2 of the RDC data was as low as in the analysis of **1**, so we calculated $^3J_{H-H}$-couplings based on a preliminary MDOC simulation to assign the measured couplings to the pro-chiral hydrogens on carbon C1 - H1A (proS) and H1B (proR) (see Fig. 5). These $^3J_{H-H}$ proved to be essential for the chiral assignment since they connect the chiral centre H2(R or S) to the prochiral hydrogen atoms.

In this process, $^3J_{H-H}$-couplings proved essential for dataset assignment from the final MDOC simulation (Table S17). The χ-probability indicated an 85 and 72% probability of assigning **2-RS** and **2-RR** to datasets 2B (Fig. 4, red bars) and 2A (Fig. 4, blue bars), respectively, improving our previous low-energy-conformer[16] and molecular-docking[19] results, with only unilateral diastereomeric discrimination for **2-SR**.

Once again, we derived conformer occurrence and populations from MDOC trajectories of the torsion angles. In **2-RR**, we observed two conformers of the five-membered ring: one principal component (94.7 %) with a maximum of 168° and one minor component (5.3 %, Fig. 5 in red) with a maximum of 36°. The torsion of the (O2–C1–C2–N1) angle was 90% in +gauche conformation with a maximum at 22° (Fig. 5). As in the case of compound **1** the pseudo-rotations of the two adjacent 5-ring systems were coordinated. The most dynamic region of both isomers of **2** was the (P–O1–C6) bridge. The rotation around the

P–O1 bond was not restricted, but the rotation around the O1–C6 bond displayed two broad maxima at approximately +100° and −100°. The angles around 0° were strongly avoided (Fig. 5). The distribution of the *P* of **2-RS** had two components: a major component (92.3 %) with a maximum at 50° and a minor component (7.7 %) with a maximum at 178°. The dihedral distributions of the **2-RS** structure mirrored those of **2-RR** (Tab. 1), but the torsion of the (O2–C1–C2–N1) angle peaked at 30° (+gauche in 96 %). When comparing **1-RR** and **2-RR** (see Tab. 1), we observed that different substituents on the phosphorus had a negligible influence on the ring conformations, but the change in configuration from **2-RR** to **2-RS** considerably affected the five-membered ring conformation.

MDOC analysis of 3

Menthol derivative **3** was prepared and separated as described in ESI. The structure was determined by NMR spectroscopy, the NMR signal assignment of two datasets 3A and 3B is in ESI. However, the absolute configuration was not assigned by XRD because compound **3** did not crystallize. For compounds **3-SR** and **3-RR**, we performed MDOC simulations using three datasets, namely 3A, 3B and 3A-ex. Dataset 3A-ex was introduced to solve the uncertainty in the assignment of the $^3J_{H-H}$-couplings. In dataset 3A-ex, the $^3J_{H-H}$-couplings of the CH$_3$

Tab. 1: Data assignment to the relative configuration and conformer analysis of compounds 1 to 3.

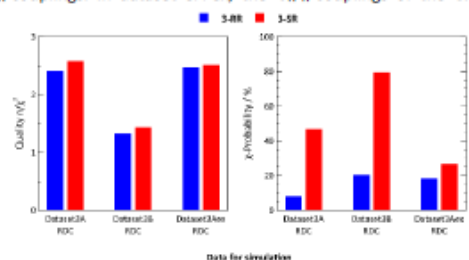| Comp. (Fig. 1) | Data | χ-Probab. [%] | Five-Membered Ring Pseudo-Rot. Angle | | | |
			Major Contrib.		Minor Contrib.	
			Max [°]	Area [%]	Max [°]	Area [%]
1-SR	1A	96	208	98	68	2
1-RR	1B	97	168	95	36	5
2-RR	2A	72	168	95	36	5
2-RS	2B	85	50	92	178	8
3-RR	3A	46	-	-	-	-
3-SR	3B	89	-	-	-	-

Fig. 6: Statistical analysis of NMR data calculated from MDOC trajectories of 3.

protons of the groups C8a and C8b to H7 were interchanged in relation to dataset 3A. The data quality of MDOC simulations of compounds **3** (Fig. 6) was nearly perfect. All RDCs were calculated within the experimental error limit, except P-C1 RDC (Tables S19 and S20). This RDC is essential to determine the configuration on the phosphorus atom because it reflects the interaction of the chiral phosphorus centre (R or S) with C1(R) of the menthol residue.

Fig. 6 displays the χ-probability of the two structures **3-RR** (blue) and **3-SR** (red) with the three datasets (3A, 3B and 3Aex). With a 79.6% probability, we can assign dataset 3B to configuration **3-SR**. However, the assignment to the other datasets, 3A and 3Aex, is inconclusive, mainly due to the outlier of the P-C1 RDCs (Tables S19 and S20). The χ-probability did not allow a final assignment. Only a preliminary assignment was performed due to the P-C1 RDC outlier: dataset 3B was assigned to the structure **3-SR**, but dataset 3A was not assigned. The highest score of **3-RR** was derived from 3Aex data set. Inspecting the P-C1 RDC, we observed that in dataset 3B the value is lower (-0.72 Hz) than the value of dataset 3A or 3Aex (-0.55 Hz). The assignment to the other datasets 3A and 3Aex is inconclusive mainly because of the P-C1 RDC outlier. If we preliminarily assign **3-SR** to dataset 3B and **3-RR** to 3A or 3Aex (3Aex was selected because of the larger χ-probability), we get the same sequence in the simulated RDC values (-0.552 < -0.469).

To try and eliminate the P-C1 RDC outlier, we performed simulations using larger weight for RDC pseudo-forces (see Table S11), albeit to no avail. Therefore, we used the same simulation parameters as in compounds **1** and **2**.

The problem was mainly caused by the rapid rotations of the menthol ring around the two bonds of the P-O1-C1 bridge (ESI, Fig. S7, like those of the P-O-C6 dihedral distribution in Fig. 5). This motion is only weakly hindered by steric effects, resulting in an averaging of the essential P-C1 RDC value due to the fast molecular motion.[6] This problem may only be solved by adding more constraints characterizing the motion of the menthol residue. Nevertheless, the MDOC results of **3** show significant improvement over the low-energy analysis which was inconclusive (Section 1.5 of the ESI).

Computational methods

MDOC simulations

To perform MDOC simulation, pseudo-energy terms are added to the energy expression of the COSMOS-NMR force field E^{ff}.[28] The pseudo-energies contain expressions depending on the differences between calculated and experimental dipolar tensors D or other scalar NMR parameters as NOE distances (denoted by R) and ^3J couplings (the different values are denoted by the indices k, l, m).

$$E = E^{ff} + \frac{k_{RDC}}{2} \sum_k \sum_{\alpha\beta} \left(D_{\alpha\beta}^{calc} - D_{\alpha\beta}^{exp} \right)^2 + \frac{k_R}{2} \sum_l \left(R_l^{calc} - R_l^{exp} \right)^2 + \frac{k_J}{2} \sum_m \left(J_m^{calc} - J_m^{exp} \right)^2$$

The first pseudo-energy term is the tensorial pseudo-energy that sums up all 9 components of the dipolar tensors (α, β run from 1 to 3). The other two terms are scalar terms that influence

structures and equilibria of conformers. The force constants k or weights of the different terms are given in table S11.

From the force-field energies, we have to calculate derivatives with respect to the coordinates of the atoms that represent the forces being the basis of the equations of motion. Since the NMR tensors depend on the orientation of the molecules within the outer magnetic the forces calculated from this term induce motions and reorientations of the molecule and parts of them. In this case, the derivatives are calculated from the rotational matrices that describe the orientations of the dipolar tensors of the molecule (for more details see Sternberg et al..[23]

As in traditional MD simulations, the temperature is controlled by a thermostat program. This is always necessary in MDOC simulations because the tensorial pseudo-forces produce heat (see Table S10). The tensorial pseudo-forces are at the begin of the simulation very large since no motional average was reached. Therefore, these pseudo-forces are gradually switched on so that their full action is reached after approximately 1 ns. Therefore, the total duration of the MDOC simulation should be much longer than 1 ns. Most parameters of the simulation were tested in former investigations but the weight and width parameters k of the pseudo forces should be adjusted in preliminary simulations (see Table S11).

The calculation of the dipolar tensors and NOE distances is straightforward[23] but for the ^3J, the Altona equations[29] were used.

Scoring calculations against experiments

The quality of MDOC simulations can be deduced using the n/χ² parameter:

$$\chi^2 = \sum_i^n \left(\frac{q_i^{exp} - q_i^{calc}}{e_i^q} \right)^2$$

$$\chi^2 = \sum_{i=1}^{n_{RDC}} \left(\chi_i^{RDC} \right)^2 + \sum_{i=1}^{n_{NOE}} \left(\chi_i^{NOE} \right)^2 + \sum_{i=1}^{n_J} \left(\chi_i^J \right)^2 + \sum_{i=1}^{n_{CS}} \left(\chi_i^{CS} \right)^2$$

In this equation, the e^q denotes the errors of the experimental properties, and a quality parameter n/χ² higher than 1.0 means that the calculated quantities q^{calc} are on average within the experimental error ranges e^q. When calculating the χ² parameter, we considered two types of data - RDC and ^3J couplings. For a good simulation, all, or at least most, summands in eq. 1 should be lower than 1.0. Under these conditions, the calculated values are within the experimental error limits, and n/χ² is higher than one.

The n/χ² value is an absolute measure of how accurately a set of calculated values represents an experiment. For comparing several datasets, a probability has been developed based on the product probabilities of deviations between calculated and experimental values. The χ-probability has been recently developed to incorporate several types of data (RDCs and ^3J couplings in this study) into a consistent tool.[26] The formalism of the χ-probability was developed on the same basis as the well established Bayesian measure DP4 developed by Smith and

Goodman[30] as stereochemical score using chemical shifts. The χ-probability differs from DP4 in the following points: (i) instead of chemical shift differences, absolute χ-values are used, and (ii) the Normal distribution is used for defining the probability that a calculated value differs from the experiment under consideration of the individual error. The χ-probability is more useful than DP4 because the error is considered and different properties can be contracted into one score.

Conclusions

To conclude, in MDOC analysis, RDC tensors can be used as constraints to simulate most features of molecular motion and to reproduce oriented NMR experiments. Discriminating relative configurations often requires introducing additional long-range constraints, such as ^3J-couplings. Using these NMR data, we can determine the relative configuration of phosphorus-containing diastereomers of rigid molecules, such as **1** and **2**. For **3**, this procedure is only partly conclusive because the rapid motion of the menthol residue leads to an outlier of the essential P-C1 RDC value. Therefore, we cannot determine the configuration of a stereogenic centre on phosphorus atom directly from P-Conly by its interaction with RDC, but using other stereogenic centres. Nevertheless, using the χ-probability, we can reach values above 85% probability of correct assignment (see Tab. 1) in most cases. These RDC-aided MDOC simulations enable us to not only identify relative configurations but also study conformational equilibria. In reactions that proceed only with selected conformers, the MDOC analysis may be used to elucidate reaction mechanisms. However, molecular flexibility still remains a challenge in RDC analysis and, therefore, new methods and improvements are highly desirable.

Author contributions

U.S. investigation, software, validation, visualization, writing, M.C.T. investigation, writing, L.T. formal analysis, A.E. synthesis, J.H. chiral separation, O.B. synthesis, and E.P. conceptualization, project administration, supervision, writing, funding acquisition.

Conflicts of interest

There are no conflicts to declare.

Data availability

The data supporting this article have been included as part of the Electronic Supplementary Information (ESI).

Acknowledgements

Notes and references

1   H. U. Blaser, B. Pugin and F. Spindler, Helv. Chim. Acta, 2021, **104**, e2000192.
2   M. Dutartre, J. Bayardon and S. Jugé, Chem. Soc. Rev., 2016, **45**, 5771-5794.
3   A. Mondal, N. O. Thiel, R. Dorel and B. L. Feringa, Nat. Catal., 2022, **5**, 10-19.
4   A. Mangoni, V. Esposito and A. Randazzo, Chem. Comm., 2003, 154-155.
5   B. Luy, K. Kobzar and H. Kessler, Angew Chem Int Edit, 2004, **43**, 1092-1094.
6   C. M. Thiele, Eur. J. Org. Chem., 2008, **2008**, 5673-5685.
7   G. Kummerlöwe and B. Luy, TrAC, Trends Anal. Chem., 2009, **28**, 483-493.
8   G. Kummerlöwe and B. Luy, Annu. Rep. NMR Spectrosc., 2009, **68**, 193-230.
9   R. M. Gschwind, Angew. Chem. Int. Ed., 2005, **44**, 4666-4668.
10 J. L. Yan and E. R. Zartler, Magn. Reson. Chem., 2005, **43**, 53-64.
11 C. M. Thiele, Concepts Magn. Reson. A, 2007, 30A, 65-80.
12 B. Böttcher and C. M. Thiele, eds., John Wiley & Sons, Ltd, 2012, DOI: 10.1002/9780470034590.emrstm1194.
13 R. Berger, J. Courtieu, R. R. Gil, C. Griesinger, M. Köck, P. Lesot, B. Luy, D. Merlet, A. Navarro-Vazquez, M. Reggelin, U. M. Reinscheid, C. M. Thiele and M. Zweckstetter, Angew. Chem. Int. Ed., 2012, **51**, 8388-8391.
14 L. Y. Liu, H. Sun, C. Griesinger and J. K. Liu, 2016, **6**, 41-48.
15 C. M. Thiele, V. Schmidts, B. Böttcher, I. Louzao, R. Berger, A. Maliniak and B. Stevensson, Angew. Chem. Int. Ed., 2009, **48**, 6708-6712.
16 M. Tichotová, A. Ešnerová, L. Tučková, L. Bednárová, I. Císařová, O. Baszczyński and E. Procházková, J. Magn. Reson., 2022, **336**, 107149.
17 F. Stanzione, I. Giangreco and J. C. Cole, Prog. Med. Chem., 2021, **60**, 273-343.
18 K. A. Peele, C. Potla Durthi, T. Srihansa, S. Krupanidhi, V. S. Ayyagari, D. J. Babu, M. Indira, A. R. Reddy and T. C. Venkateswarulu, Inform. Med. Unlocked, 2020, **19**, 100345.
19 M. C. Tichotová, L. Tučková, H. Kocek, A. Růžička, M. Straka and E. Procházková, Phys. Chem. Chem. Phys., 2024, **26**, 2016-2024.
20 M. Di Pietro, P. Tzvetkova, T. Gloge, U. Sternberg and B. Luy, Liq. Cryst., 2020, **47**, 1-15.
21 P. Tzvetkova, U. Sternberg, T. Gloge, A. Vazquez and B. Luy, Chem. Sci., 2019, **10**, 8774-8791.
22 M. Di Pietro, U. Sternberg and B. Luy, J. Phys. Chem. B, 2019, **123**, 8480–8491.
23 U. Sternberg, P. Tzvetkova and C. Muhle-Goll, Phys. Chem. Chem. Phys., 2020, 22, 17375-17384.
24 U. Sternberg and R. Witter, Molecules, 2022, 27, 7987.
25 J. G. P. Wicker and R. I. Cooper, J. Chem. Inf. Model., 2016, 56, 2347-2352.
26 U. Sternberg and C. Farès, Phys. Chem. Chem. Phys. , 2022, 24, 9608-9618.
27 C. Altona and M. Sundaralingam, J. Am. Chem. Soc., 1972, 94, 8205-8212.
28 M. Möllhoff and U. Sternberg, J. Mol. Model., 2001, 7, 90-102.

29   C. A. G. Haasnoot, F. A. A. M. De Leeuw, H. P. M. De Leeuw   30   S. G. Smith and J. M. Goodman, J. Am. Chem. Soc., 2010,
        and C. Altona, Biopolymers, 1981, 20, 1211-1245.   132, 12946-12959.