BIOMOLECULAR PHOTOSWITCHING WITH THE AZOBENZENE CHROMOPHORE AND LOGP DETERMINATION OF AZO DYES

by

Christopher T. Brown, B.S.

A thesis submitted to the Graduate Council of
Texas State University in partial fulfillment
of the requirements for the degree of
Master of Science
with a Major in Chemistry
May 2017

Commitee Members:

William J. Brittain, Chair

Alexander Kornienko

Steven Whitten

FAIR USE AND AUTHOR'S PERMISSION STATEMENT

Fair Use

Duplication Permission

ACKNOWLEDGEMENTS

The author would like to acknowledge the three committee members William J. Brittain, Steven Whitten, and Alexander Kornienko for their guidance and supervision of this work. I would also like to thank Dr. Ben Shoulders for providing advanced training in nuclear magnetic resonance spectroscopy, Dr. Kevin Lewis for providing access to the Varian Cary 400 spectrophotometer, as well as Dr. Shiva K. Rastogi for training and support with synthesis and purification of any materials used in the study.

Along with the aforementioned individuals, the author would like to recognize the contributions of former and current lab members Joseph Lamas, Cindy Salinas-Hernandez, Katherine Martin, Scott Barrett, Steve Gralinski, and Hailey Anderson for their technical assistance, and contributions to experimental and theoretical approaches to the material presented herein.

TABLE OF CONTENTS

Page

LIST OF TABLES

Table **Page**

LIST OF FIGURES

Figure **Page**

LIST OF SCHEMA

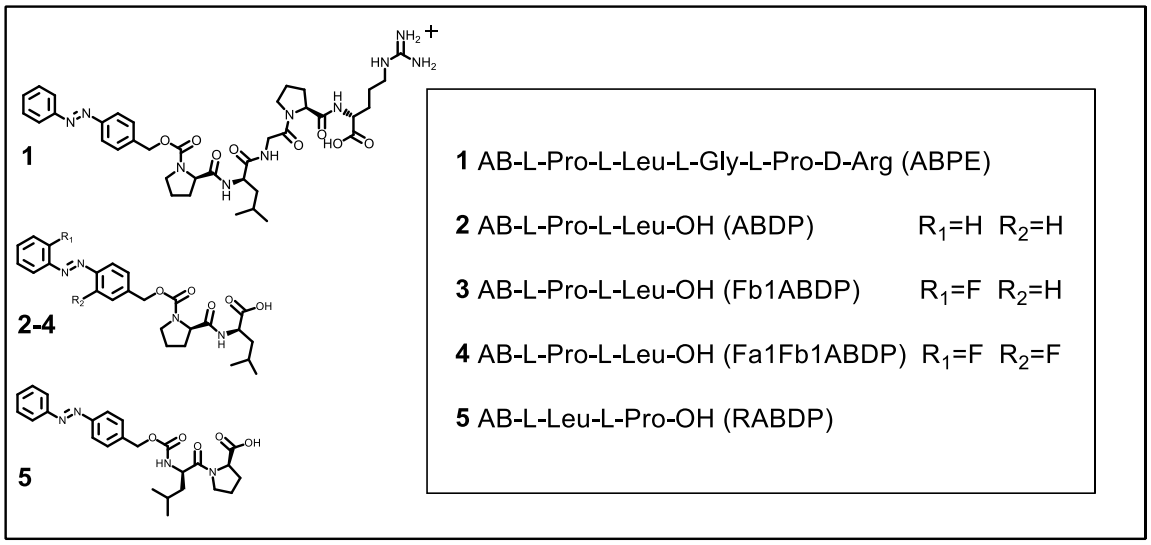
Scheme **Page**

I. INTRODUCTION AND BACKGROUND

Understanding light induced sensing and response at the molecular level provides opportunities for the creation and application of new photo-tunable materials[1]. The azobenzene chromophore in particular is capable of directing the behavior of a molecular system using light,[2,3] with the advantage of being noninvasive, reversible,[4] and highly specific. While these molecules have been well studied, there remains a lack of fundamental information regarding the behavior of poorly soluble azobenzenes in water.[5] When there is limited solubility, saturated aqueous solutions of azobenzene show unusual ratios of *cis* and *trans* isomers. Many of these compounds have also demonstrated enhanced thermal stability of the *cis* photoisomer in aqueous solvent.[6] While a number of studies have exploited differential solubility between *cis* and *trans* isomers on functionalized surfaces, investigations in photopharmacology to enhance membrane permeability of drugs and medicinal agents has been only marginally explored.[7,8] In spite of the great deal of information about azobenzenes in general, the molar aqueous solubility for both isomers remains unanswered for the majority of these compounds.

Quantifying aqueous solubility measurements for separated isomers by dissolution, and filtering is inherently problematic due to the very limited intrinsic solubility of many of these compounds. Solvation of several milligrams of material can require 2 or more liters of aqueous solvent, and care must be taken to filter out any particulates or colloids that may have formed. Measuring *cis* and *trans* isomers separately is further complicated by thermal relaxation of the *cis* isomer to *trans* and the

dynamic equilibria that exists between them. Lastly, bioconjugates with DNA and peptides, or any derivatives containing ionizable or highly polar functional groups cannot be separated by normal-phase flash chromatography. Due to the challenges presented, a method has been developed which correlates experimentally derived retention times (tR) from reverse-phase high performance liquid chromatography (RP-HPLC) to LogP values. This method, which is based off of Abraham solvation theory, has the advantage of rapid determination of LogP values for both *cis* and *trans* conformers,[9] and is compatible with azobenzene bioconjugates. Although LogP values are indirect assessments of solubility, partition coefficients are able to provide information about the behavior of these molecules at interfaces and in biological systems.[10]



Scheme 1. Molecular structures of azobenzene bioconjugates purchased(**1**) and synthesized(**2-5**) for solubility and thermal stability studies.

The design of biocompatible azobenzenes with stable photoswitching is of particular importance for *in vivo* applications. Prolonged thermal half-lives of the *cis* isomer for an azobenzene functionalized dipeptide and pentapeptide was previously reported by Chambers and Haworth.[6] Originally used as a collagenase assay,[11] the

pentapeptide substrate with sequence Azo-L-Pro-L-Leu-L-Gly-L-Pro-D-Arg (ABPE) is cleaved at the Leu-Gly peptide bond and the Azo-Pro-Leu dipeptide (ABDP) is formed as a product (**Scheme 1**). Solutions of ABPE in D_2O were observed to be comprised of 30% *cis* azobenzene after 15 days under dark conditions.[6] Following this observation, azobenzene-dipeptide derivatives were designed with features including an Azo-Leu-Pro sequence reversal (RABDP), and the original Azo-Pro-Leu sequence with fluorine substituents placed *ortho* to the N=N double bond (Fb1ABDP, Fa1Fb1ABDP). According to Chambers and Haworth,[6] it was believed that interactions between the leucine sidechain and aromatics provided additional free energy to the *cis* conformation through hydrophobic interactions. From this conjecture, it was anticipated that sequence order of the peptide substituents would have a significant effect on thermal half-lives. Also worth mentioning is the effect of σ-electron withdrawing groups (EWGs) positioned *ortho* to the azo bond, which stabilizes the *cis* isomer by lowering the energy of the non-bonding orbital.[12]

The photochemical behavior of azobenzene has been extensively studied via UV-Vis absorption spectroscopy and Nuclear Magnetic Resonance Spectroscopy (NMR), both of which are capable of generating distinct spectra for the *cis* and *trans* photoisomers.[13,14] Spectroscopic techniques are important for understanding dynamic *cis*/*trans* equilibria for azo-dyes under specific conditions. UV-Vis absorption spectroscopy has the sensitivity to detect azobenzene species at micromolar concentrations although spectral features often overlap between species and it is not capable of accurately resolving mixtures. NMR spectroscopy will distinguish structural features between *cis* and *trans* isomers as well as different azobenzene species, and can

be used for near quantitative measurement of a mixture of species in solution. These tools provide for a way to probe and quantify the dynamic behavior of azobenzene's photochemistry in biologically relevant aqueous solutions.

The difference in solubility between the *cis* and *trans* isomers in azo dyes is of particular interest for the purpose of creating materials with tunable properties. Furthermore, differential solubility affects the pharmacokinetics of these molecules and the ways in which they will be absorbed and distributed throughout a biological system. Poorly soluble molecules have higher lipid solubility, or lipophilicity.[15] A molecule with high lipophilicity will generally accumulate within the lipid portion of the cell membrane. On the other hand, molecules with high aqueous solubility have low lipophilicity and are generally incapable of penetrating the cell membrane.[16] Compounds with low lipophilicity tend to exhibit poor oral absorption through the intestinal membrane. In general, molecules with ideal absorption, distribution, metabolism, and excretion (ADME) are typically those with moderate lipophilicity, because they are capable of penetrating lipid membranes and tissues and more readily move through both aqueous and lipid phases. Lipophilicity is important when considering the route of administration for a drug, or whether or not adjuvants must be employed to aid ADME properties of a pharmacological agent.[17] The highly reversible *cis*/*trans* isomerization along with thermal stability and prolonged half-lives of certain azobenzenes provides a non-invasive means for modulating ADME properties through photopharmacology. Little, if any, information has been reported about the lipophilicity of azo dyes, and the literature does not distinguish nor report the lipophilicity of both the *trans* and *cis* isomers. Standard lipophilicity measurements are carried out by dissolving a compound in a mixture of

octanol and water and then determining the partition coefficient between two immiscible phases by separating the phases, removing solvent, and weighing the material obtained from each phase.[15] For azo dyes and azobenzene-peptide conjugates, the limited solubility, and small amounts of available material makes the traditional method error prone, inefficient and time-consuming. Instead, the partition coefficient of sparingly soluble azobenzenes was measured according to a previously developed method.[18]

The work presented herein has two main objectives; a study of the thermal stability and prolonged half-lives of azobenzene-peptide bioconjugates, and determination of lipophilicity for a family of azo dyes. A molecule has been prepared where the substituent order of the dipeptide mentioned in the Haworth paper has been reversed, with the purpose of examining the effect of amino acid sequence on the stability of *cis* conformers. Two derivatives of the original sequence have also been prepared with *ortho* fluorine substituents. The photoisomers of these molecules were structurally characterized via NMR spectroscopy. Kinetics experiments were carried out by UV-Vis absorbance spectroscopy in aqueous solution. Solubility studies for azo-dyes were quantified via separation of isomers through flash chromatography and absorbance spectroscopy. Lastly, lipophilicity of a family of azo dyes and azobenzene-peptide conjugates was measured using RP-HPLC.

II. BIOMOLECULAR PHOTOSWITCHING WITH AZOBENZENE

2.1 ^{1}H NMR Spectroscopy of Azobenzene-Peptide Conjugates

The azobenzene chromophore exists as two distinct structural isomers that display different absorption spectra. The chemical shift of an NMR signal is largely dependent upon the shielding effect conferred by electrons surrounding an NMR active nucleus. The molecular features present in the *cis* and *trans* isomers are such that the signals for the protons contained in each isomer are highly distinguishable. NMR spectroscopy is used to observe azobenzene photodynamics, and is capable of probing changes in a molecule's local environment. Differences in molecular structure between *cis* and *trans* isomers of azobenzene-peptide conjugates were determined with NMR spectroscopy by observing and comparing samples irradiated with ultraviolet and visible light. The rates of isomerization were also measured and correlated to kinetics carried out with UV-vis absorbance described in a subsequent chapter.

2.2 Experimental Methods for ^{1}H NMR Spectroscopy of Azobenzene Bioconjugates

The azobenzene-dipeptide used in NMR analysis, (S)-4-methyl-2-((S)-1-(((4-((E)-phenyldiazenyl)benzyl)oxy)carbonyl)pyrrolidine-2-carboxamido)pentanoic acid (ABDP), was purchased from a commercial source (Bachem and Sigma-Aldrich, USA). All reagents and solvents were purchased from commercial sources (Acros Organics and Sigma-Aldrich, USA) and used without purification. NMR spectra were recorded on a Bruker Avance III 400 and Bruker Ascend 500 spectrophotometer. ^{1}H and ^{13}C NMR shifts are reported relative to TMS and ^{19}F shifts are relative to C_6F_6 (-164.9 ppm).

(400.13 MHz for ^{1}H; 100.61 MHz for ^{13}C) Chemical shifts (δ) are reported in ppm relative to the TMS internal standard.

For ^{1}H NMR all samples were prepared by measuring 1-2mg of material in a glass vial and adding 1mL of deuterium oxide. Due to the limited solubility of these compounds, samples were sonicated for 1 hour at 35C before filtering with GHP acrodisc 25µm syringe filters into dry NMR tubes. After dissolution, sonication, and filtering, samples were allowed 1 hour to equilibrate to room temperature prior to spectral acquisition. Dark photostationary states (PSS) were obtained by keeping samples away from light for at least one week prior to NMR acquisition. Photoirradiation was carried out via a XeHg lamp and UV bandpass filters described in detail in the subsequent chapter. Samples were irradiated for a total of 30 minutes in quartz cuvettes under constant stirring. 1-Dimensional (1D) proton NMR spectra were collected without spinning and an acquisition time of 3s, delay time at 5s, NS=64 and DS=4. 2D COSY spectra were acquired as follows; F2 FID size, 2048, F1 FID size 128, NS 1, DS 8, AQ 0.25 sec, P1 2 usec.

2.3 ^{1}H NMR Spectral Assignments for ABDP

Assignments of ^{1}H NMR spectra for azobenzene-peptide conjugates are necessary in order to confirm the presence of the molecule of interest and identify *cis* and *trans* photoisomers. After the assignment of proton resonances, the composition of *cis* and *trans* isomers present at the dark and UV irradiated photostationary states were approximated by integrating the area under aromatic resonances assigned for *cis* and *trans* isomers and calculating a percent composition. The following procedure for

azobenzene-peptide conjugates in this study is demonstrated using 4-Phenylazobenzyloxycarbonyl-Pro-Leu-OH (ABDP)

The aromatic portion of ABDP's spectra has 5 sets of protons that exist in distinct chemical environments. For the *trans* isomer (**Figure 1, top)**, signals **b** and **c** appear at lowest field as two sets overlapping doublets, integrating to four. Signals **a**, **d**, and **e** appear at higher field as a group of two doublets and a triplet integrating to five. Signals for the *cis* isomer appear at higher field due to increased shielding from a lack of conjugated pi bonds. The set of two overlapping doublets and a triplet at δ7.3 - 7.39 ppm were assigned to *cis* protons **a'**, **d'**, and **e'**. The lone pair on each nitrogen is capable of donating electrons into each phenyl ring through resonance, primarily affecting the chemical shift of *ortho* protons. This is observed as a group of overlapping doublets at δ7.0 assigned to protons **b** and **c**.

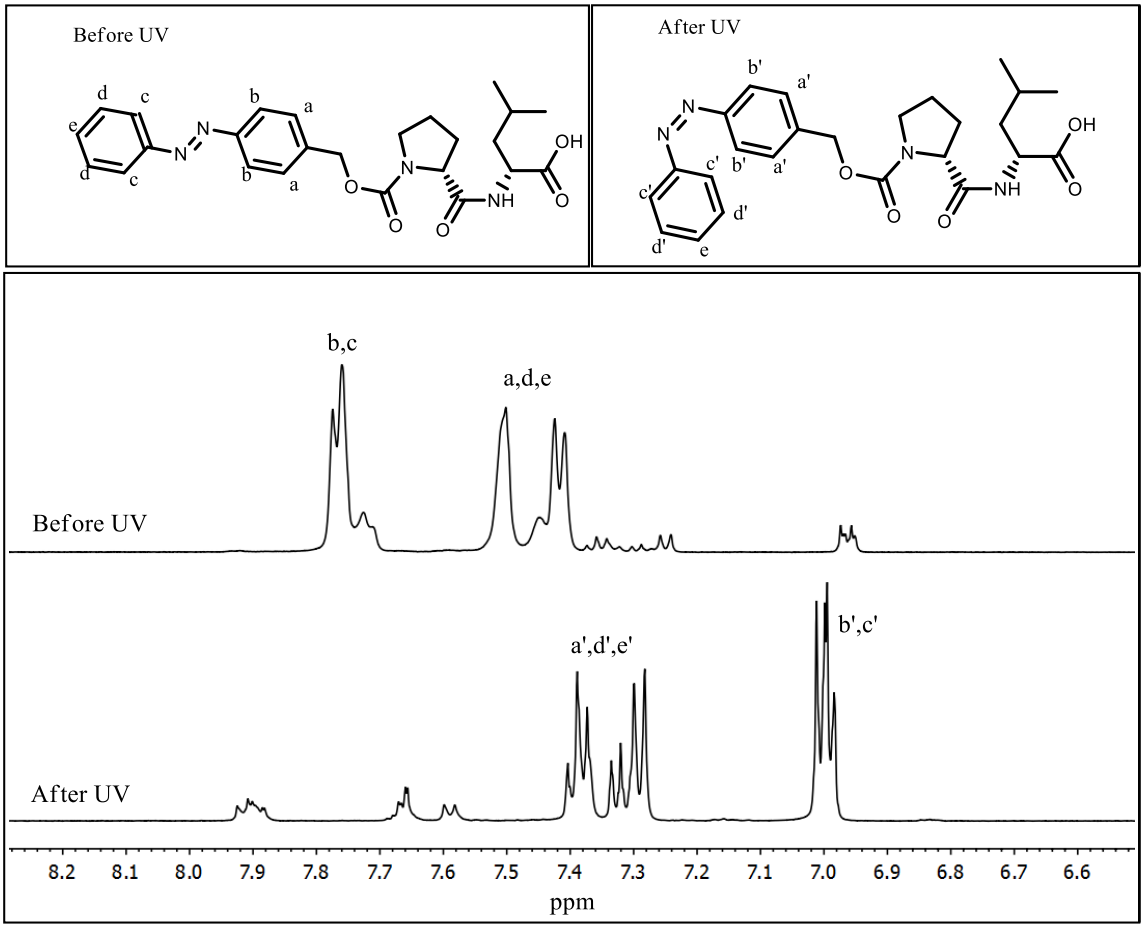


Figure 1. 500 MHz 1D Proton NMR Spectra overlay of photostationary states (PSS) before UV irradiation and after UV irradiation. Assignments for aromatic resonances are as follows; Trans Azobenzene (top) δ7.95 b,c δ7.65 d,e δ7.6 a. Cis Azobenzene (bottom) δ7.39 e', δ7.3 a',d' δ7.0 b',c'.

Structural comparison of the entire NMR spectra for the *cis* and *trans* azo-dipeptide were made using a 500 MHz Bruker Spectrometer (**Figure 2**). The dark PSS of ABDP was obtained by preparing a sample in an NMR tube as previously described and keeping the tube away from light for a period of 4 weeks. After spectral acquisition, the sample was irradiated with UV light for a period of 30 min with 5 min intervals of shaking. Next, a spectra of UV-irradiated ABDP was acquired and the two were compared for structural changes. Acquisitions used a pulse-program with pre-saturation

water suppression and calibrated to δ4.79 at the water signal. Spectral acquisition parameters were as follows; DS 4, NS 64, FID 32k, d1 1sec, LB 0.

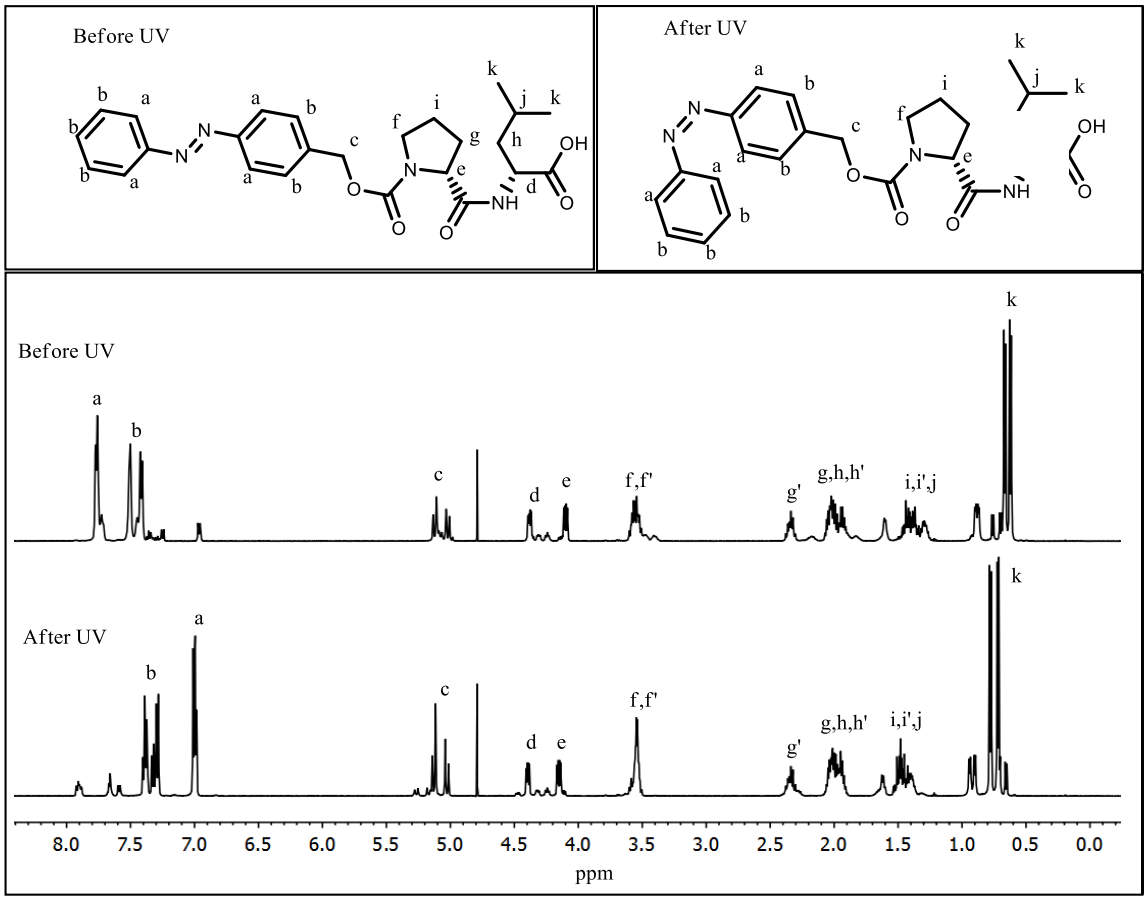


Figure 2. 500MHz ^{1}H NMR spectra comparison of ABDP in dark PSS (top) and UV-irradiated PSS (bottom). Structure of ABDP with spectral assignments is pictured above.

The dark adjusted spectra of ABDP had a mixture of *cis* and *trans* ABDP with a predominant *trans* configuration as observed in the aromatic region, with δ7.6 and δ7.4 assigned to 4 aromatic protons designated **a**, 5 aromatic protons labelled **b**. The benzyllic protons, **c**, are diasterotopic and show up as two doublets at δ5.1 and δ5.0. The methine α protons (Hα), **d** and **e**, appear at δ4.4 and δ4.1. The leucine Hα, **d**, is more deshielded compared to the to the proline Hα, **e**. This effect can be explained by **d** being located

between a carboxyllic acid and an amide, where **e** is located between two amides. At δ3.5 is a multiplet integrating to 2 that is assigned to **f,f'**. These 2 protons are diastereotopic and each individual proton would be split by its geminal and 2 vicinal neighbors. The chemical shift of **f,f'** agrees with the patterns of chemical shift for N-methyl-2-Pyrrolidone,[19] where the ring protons adjacent to the nitrogen are the most deshielded. The positioning of **f,f'** γ from the nearby carbonyl oxygen also further adds to this effect. The γ effect is where electron density surrounding a proton is redistributed towards its vicinal carbon due to electron repulsion from oxygen atoms positioned 3 bond lengths away.[20] This effect also explains the signals at δ2.3 and δ2.0, where the first multiplet labelled **g'** integrates to 1 and is shifted δ0.3 ppm from its neighboring multiplet assigned to **g,h,h'**. Assignments for **i,i'** appear at δ1.5 which is consistent with proton shifts of cyclic amides discussed previously.[19] The signal at δ1.3 is a multiplet integrating to 1, agrees with the chemical shift for a tertiary carbon and is assigned to the methine proton **j**. The two methyl groups, **k**, appear at δ0.7 ppm as a doublet of doublets integrating to 6.

Some important observations to point out in are the differences in chemical shifts between *cis* and *trans* azobenzene for certain signals, most notably the leucine methyl groups, **k**, seen in detail in **Figure 3**. The top spectra also shows weaker signals at δ3.4, δ2.2, δ1.8, δ1.4 ppm likely to be the result of rotational isomers for *cis* peptide bonds.[21] Shown on the bottom of **Figure 3**, these weaker signals are observed to shift an average of δ0.1 ppm downfield upon photoisomerization of azobenzene to *cis*, suggesting changes to the local environment of those protons.

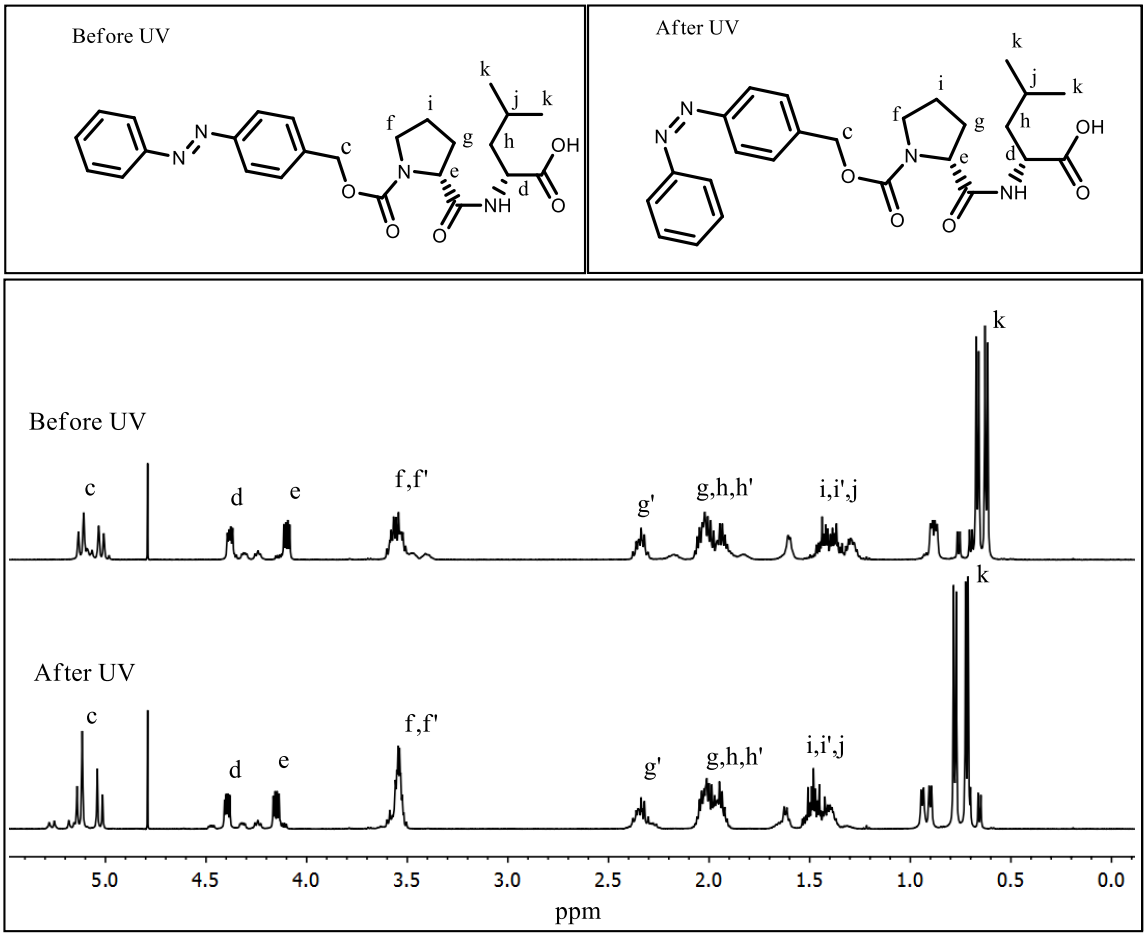


Figure 3. 500MHz ^{1}H NMR spectra comparison of aliphatic region for ABDP in dark PSS (top) and UV-irradiated PSS (bottom). Structure of ABDP with spectral assignments is shown in Figure 3.

Structural elucidation of complex molecules such as ABDP can be problematic using 1 dimensional methods due to the presence of cyclic alkanes, branched alkanes, amides, and carbonyls with anisotropic effects. 2 dimensional ^{1}H-^{1}H correlation spectroscopy (COSY) in **Figure 4** was used to confirm connectivity of resonance assignments. There were only 3 correlation signals observed, likely due to the limited solubility of ABDP in water. Signal **e** and **f,f'** are observed to couple to signal **i**, where variable amounts of through-bond coupling can be explained by the influence of dihedral

angle on coupling constants.[22] Lastly, coupling is confirmed between signal **k** and **j** at δ0.6 and δ1.2, respectively.

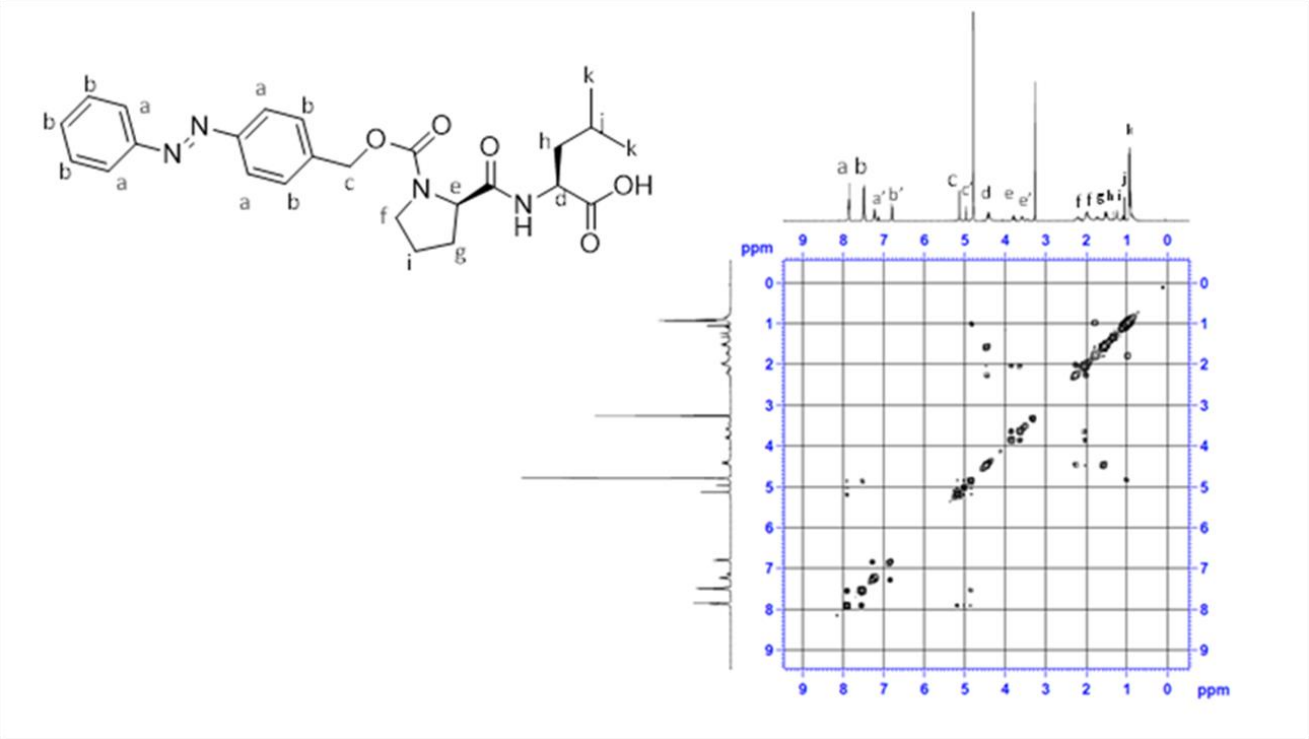


Figure 4. 400 MHz 2D COSY spectra of dark PSS for ABDP aliphatic region.

2.4 Effect of Solvent and Added Surfactant on *Cis/Trans* Ratios

There is a significant contribution of the *cis* azobenzene moiety for ABDP when kept in dark conditions as indicated by the presence of signals in the sample kept in dark. When ABDP is irradiated with UV light, the *trans* signals decrease while the *cis* signals increase. In order to further investigate the stability of the *cis* moiety in water, the dark PSS for ABDP was observed and compared between deuterium oxide, deuterium oxide with 100mM sodium dodecyl sulfate, and deuterated methanol (D_2O, D_2O + 100mM SDS, MeOD) using ^1H NMR spectroscopy.

The amount of surfactant required for stable micellar aggregation, otherwise known as the critical micelle concentration (CMC), for SDS occurs at or above 8mM concentrations[23]. A 100mM solution of SDS in deuterium oxide was used to investigate via ^{1}H NMR the effect of added surfactant on aqueous *cis/trans* ratios for azobenzene bioconjugates.

$$[M] = \frac{[S]-CMC}{N} \qquad \textbf{Equation 1}$$

The mean aggregation number, N, for SDS micelles at 25C in water has been experimentally determined to be at or near 62 SDS monomers per micelle[24]. The concentration of micelles in solution, [M], would be equal to the difference between the concentration of surfactant, [S] and the CMC, divided by the mean aggregation number. According to the aforementioned formula, a 100mM solution of SDS would have a micellar concentration of 1.5 mM, and the number of micelles to molecules of ABDP at the measured concentrations (0.1mM) would be 15 to 1.

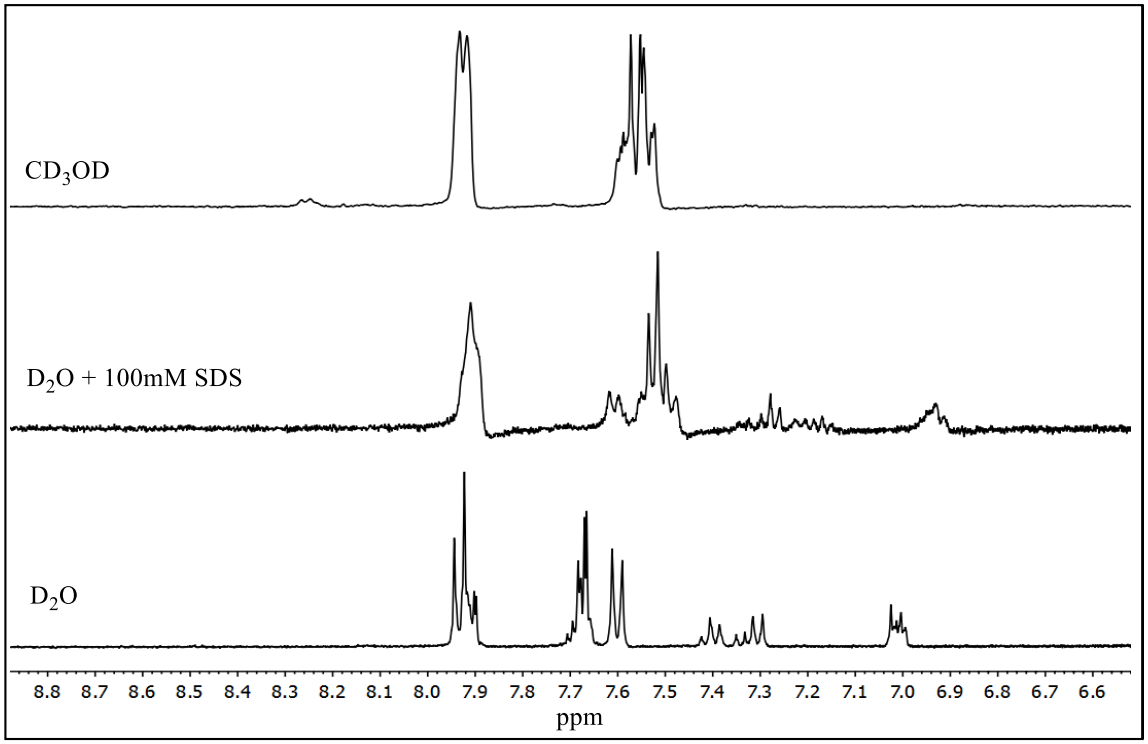


Figure 5. 400 MHz Proton NMR spectra overlay of Dark PSS for ABDP in MeOD (top), D_2O + 100mM SDS (middle), D_2O (bottom).

In addition to Deuterium Oxide, and 100mM SDS solutions, deuterated methanol (MeOD) was selected because it is a polar solvent with similar characteristics to water. In **Figure 5**, The dark PSS of ABDP in MeOD had two distinct *trans* signals integrating to four and five respectively, with nearly zero contribution of *cis* aromatic signals. In 100mM of SDS and deuterium oxide, the lower field *trans* signal has significant line broadening in comparison to ABDP in D_2O. The higher field *trans* signals also have significant chemical shift differences in comparison to the D_2O counterpart. The signals for the *cis* isomer are reduced in comparison to ABDP in D_2O, and there is also significant line broadening in the signal at δ7.9 ppm.

2.5 Photodynamics of Azobenzene Peptide Conjugates

At the atomic level, energy is quantized and a chemical species can only exist in discrete, defined energy states. This quantization of energy results in individual species absorbing or emitting electromagnetic radiation at specific frequencies. Azobenzene behaves such that, upon irradiation with certain frequencies of light, the molecule undergoes photoisomerization.

In many azo dyes, the *cis* and *trans* isomers absorb light at different wavelengths, and therefore the rates of isomerization can be determined by monitoring the change in absorbance of these wavelengths. This method is quantitative due to the direct relationship between absorbance and concentration as defined by Beer-Lambert Law;

$$A = \varepsilon \cdot l \cdot c \qquad \textbf{Equation 2}$$

Where absorbance, A is the product of the molar extinction coefficient ε, the path length of a given sample l, and the concentration c. The extinction coefficient ε is a parameter with units of $M^{-1}\ cm^{-1}$ that describes the ability of a particular molecule to absorb light at a specific wavelength.

$$[cis] \xrightarrow{heat} [trans]$$

Isomerization from *cis* to *trans* azobenzene is an intramolecular first-order process and occurs under the absence of light source via a thermal mechanism.

$$\mathrm{v} = \mathrm{k}[cis] \qquad \textbf{Equation 3}$$

The rate law of this reaction is described where the velocity of the reaction, v, is equal to the rate, k, with units of inverse seconds (s^{-1}), multiplied by the concentration of

reactant, [*cis*]. Similarly, the velocity of the reaction can be described as being equal to the dissapearance of [*cis*] azobenzene;

$$v = -\frac{d[cis]}{dt} \qquad \textbf{Equation 4}$$

Setting these expressions equal to each other, separating terms on each side, and integrating each term yields the first order integrated rate law.

$$\ln[cis] = -kt + \ln[cis]_0 \qquad \textbf{Equation 5}$$

Expressed in linear form, the first order integrated rate law states that the natural log of the concentration of [*cis*] present is a function of time, *t*, given that the rate, *k*, and initial concentration $[cis]_0$ is known. Using Beer-Lambert law as previously described, and assuming that the path length, l, and molar absorbtivity, ϵ, remain constant, the concentration of [*cis*] present will be directly related to it's measured absorbance. The rates of thermal isomerization were determined for azobenzene dipeptide conjugates with UV-vis absorption spectroscopy at 25°C and 50°C in aqueous solution.

2.6 Experimental Methods for Determination of Rate Constants for Thermal *Cis-Trans* Isomerization

Spectral acquisitions were conducted using an Ocean Optics HR2000+ CCD array detector, QPOD temperature controlled cuvette holder with stirring and a DH2000 deuterium/halogen light source. All components were connected by 600 μm diameter quartz fiber optic cables with SMA905 connectors. Spectra are reported for the 250-650 nm spectral region. The detector configuration is as follows; 25 μm slit aperture, no order

sorting filter, no detector collection lens, grating is H1 at 300nm. Collection parameters are as follows; integration time of 6 msec, 20 scan averaging, and a boxcar width of 3.

Sample irradiation was performed using a XeHg Lamp and Newport Power Supply set to a constant power of 250 Watts. Bandpass filters were used to select UV or visible wavelength ranges for irradiation (**Figure 6**), and the irradiation beam was positioned orthogonally to the probe beam for all spectral acquisitions. The sample volume in all spectroscopic measurements was ~2mL, and a 1cm quartz cuvette with a septum cap was used for all samples. The photostationary state (PSS) were recorded by irradiating the sample continuously with UV light until the absorbance of the π-π^* and n-π^* bands exhibited no further change with continued irradiation.

Ultraviolet-visible kinetics acquisitions were collected using a Varian Cary 400 spectrophotometer with a dual peltier temperature controlled cuvette holder and stirrer. All sample irradiations were performed and monitored via the irradiation-probe setup previously described. After reaching the photostationary state, the cuvette was immediately transferred to the Varian Cary 400 spectrophotometer. Scans were taken every 10 minutes for a period of 1 hour, and subsequent scans were taken every hour for a period of 71 hours.

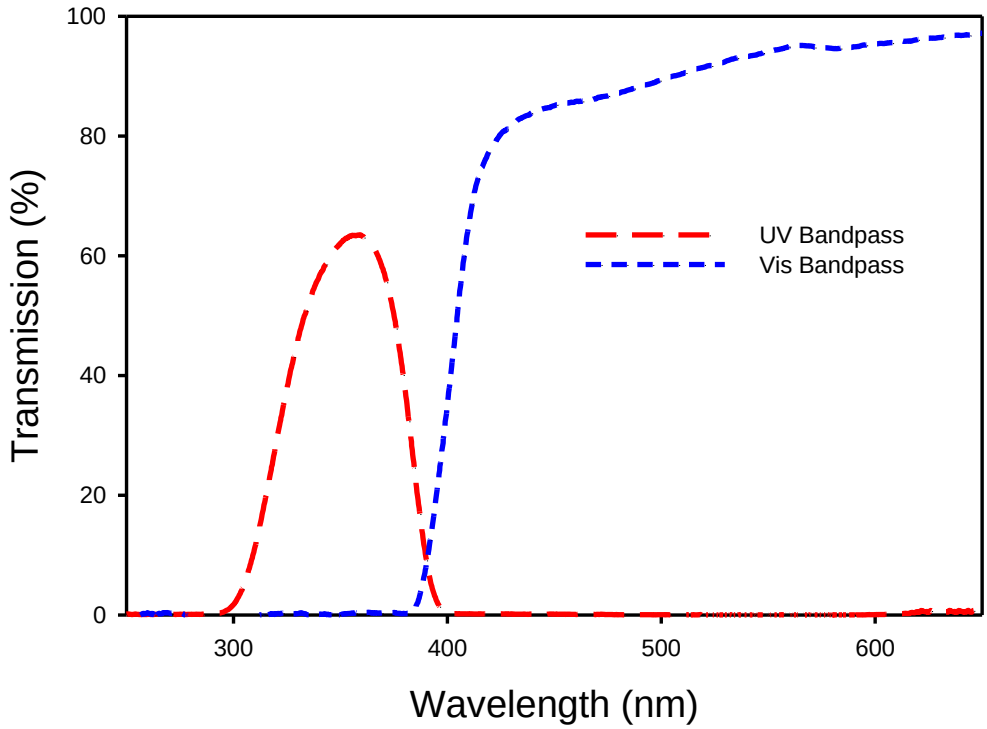


Figure 6. Transmission spectra of bandpass filters used for irradiation experiments.

2.7 Rates of Thermal *Cis-Trans* Isomerization for Azo-Peptide Conjugates

Plots of the thermal *cis-trans* isomerizzation of azobenzene bioconjugates were made by taking the natural log of the n-π* absorbance maximum at given time intervals. The observed maxima is solvent dependent, with absorbance shifts depending on solvent polarity.[25] Peak maxima used in kinetics experiments are shown in **Table 1**. The absorption spectra of ABDP and related compounds from 250-650nm has two absorbance maxima (**Figure 7**) corresponding to two electronic transitions. The higher energy π-π* transition is the result of promotion of an electron in a bonding π orbital to an anti-bonding orbital. The peak maxima for the π-π* transition are different between the *cis* and *trans* isomers, ranging between 320-325nm for the *trans* isomer, and 290-295 for the *cis* isomer.

Table 1. Absorbance maxima (λ_{max}) for ABDP in water, methanol, and SDS micelles.

Solvent	$\lambda_{\pi\text{-}\pi^*}$, trans	$\lambda_{n\text{-}\pi^*}$, trans	$\lambda_{\pi\text{-}\pi^*}$, cis	$\lambda_{n\text{-}\pi^*}$, cis
Water	325	427	293	428
Methanol	321	433	290	432
SDS Micelles	324	431	293	430

The lower energy n-π* transition results from promotion of an electron in a non-bonding orbital to a π anti-bonding orbital and occurs near 425nm. Solvent has a small, but noted effect on peak maxima of azobenzene bioconjugates. Moving from water to methanol, the λ_{max} for the π-π* transition undergoes a hypsochromic blue-shift of 4 nm for *trans* and 3nm for *cis*. The opposite is observed for the n-π* transition, with a bathochromic red-shift of 6nm for *trans* and 4nm for *cis*. There was a small shift in absorbance maxima between water and SDS micelles for the π-π* transition of *trans* ABDP and n-π* transition for *cis* ABDP of 1nm and 2nm, respectively.

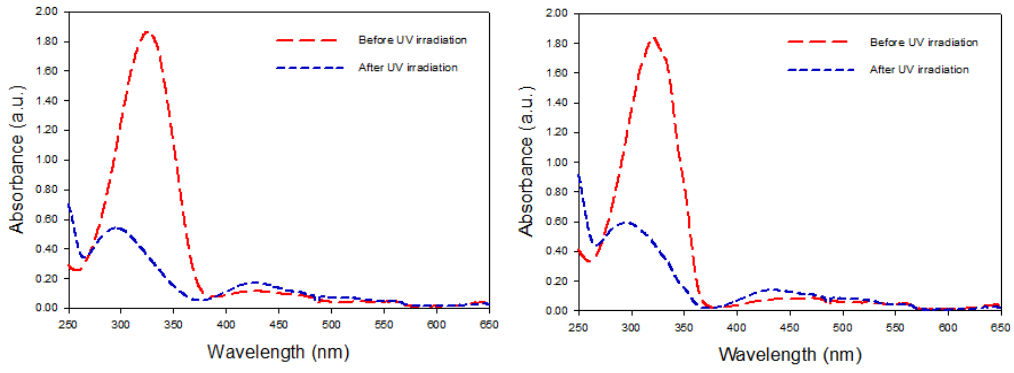


Figure 7. UV-vis absorption spectra of ABDP in water (left) and methanol (right), showing absorption before and after irradiation with UV light.

The π-π* absorbance for the *trans* isomer in water has a λ_{max} at 325nm, while the *cis* isomer has a λ_{max} at 293nm. The n-π* transition was chosen to monitor the dissapearance of *cis* for thermal isomerization kinetics. At 50°C, the kinetics plots shown in **Figure 8** of ABDP show a linear trend in both methanol and water. A comparison of the rate constants in **Table 2** for ABDP, the reverse dipeptide, RABDP, ABDP in SDS micelles, and fluorinated analogs show negligible differences in the rates of isomerization. There is a trend of increasing rates of isomerization going from polar to non-polar solvents. Fluorinated azobenzene-dipeptides show a trend of decreasing rates of isomerization as the number of fluorine substituents increases.

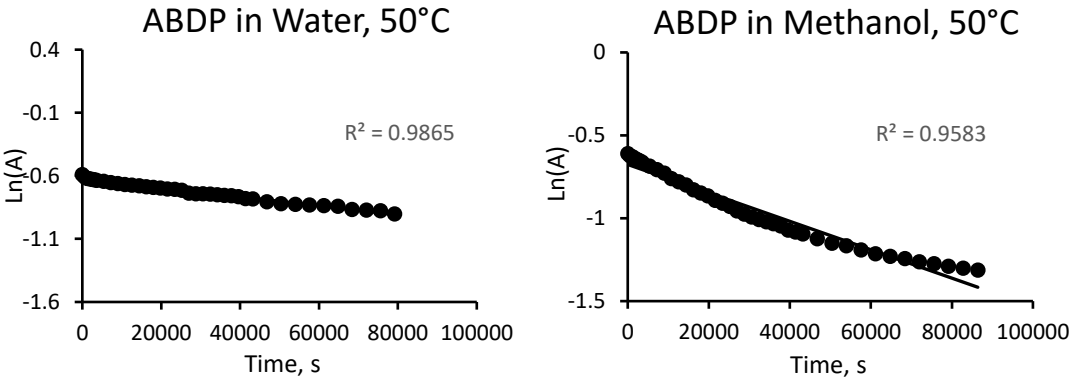


Figure 8. Kinetics plots of ABDP in water (left) and methanol (right), 50°C.

Table 2. Thermal isomerization rates for azobenzene-dipeptide conjugates in methanol (right column) and water (left column). See appendix for compound list.

name	k x 10^{-5}·s^{-1} at 50°C, Methanol	k x 10^{-5}·s^{-1} at 50°C, Water
azobenzene	2.6 ± 0.07	-
ABDP	0.87 ± 0.03	0.36 ± 0.01
ABDP + SDS micelles	-	0.60 ± 0.03
RABDP	0.85 ± 0.03	0.35 ± 0.01
Fb1ABDP	0.51 ± 0.01	0.17 ± 0.02
Fb1Fa1ABDP	0.27 ± 0.01	-

Half-lives of the thermal isomerization from *cis* to *trans* were calculated by rearranging the first-order rate law as follows;

$$\frac{ln\frac{1}{2}[cis]_0}{ln[cis]_0} = -kt_{1/2} \qquad \textbf{Equation 5}$$

where the natural log of 1/2 the initial concentration of *cis* divided by the initial concentration of *cis*, is related to the rate constant, k, and half life, $t_{1/2}$.

$$\frac{\ln 2}{k} = t_{1/2} \qquad \textbf{Equation 6}$$

By the properties of logarithms, this equation simplifies to half life, $t_{1/2,}$ being equal to the natural log of 2 divided by rate constant, k. Shown in **Table 3**, the half life of ABDP in water is slightly greater than double the half-life in methanol. Dipeptide sequence order has a negligible difference on half-life, with azo-pro-leu (ABDP), and azo-leu-pro (RABDP) having nearly identical $t_{1/2}$ values.

Considering the stable *cis* moeity observed via NMR, these results suggest that the molecular structure of azobenzene-peptide conjugates is not responsible for this phenomena. It is much more likely that the *cis-trans* equilibria for the dark PSS is altered due to the combined effect of limited solubility in water, and differential solubility between *cis* and *trans* azobenzene.

Table 3. Half lives of azobenzene-dipeptide conjugates in methanol (left column) and water (right column) at 50°C.

name	$\tau_{1/2}$, Methanol at 50°C	$\tau_{1/2}$, Water at 50°C	$\tau_{1/2}$, DMSO at 35°C
Azobenzene	8h	-	17h
ABET	-	-	35h
ABDP	22h	53h	-
RABDP	23h	55h	-
Fb1ABDP	38h	113h	-
Fb1Fa1ABDP	71h	-	-

One of the distinguishing features of the dipeptide substituted azobenzenes is a modest increase in the thermal stability of the *cis* isomer in methanol at 50°C. This observed increased in stability of the *cis* isomer is likely the result of the carbamate group acting as an electron donating group (EDG) into the conjugated ring system. The electronic contribution of the carbamate linker would direct more electron density to the N=N substituent and increase the energetic barrier for isomerization. Similar increases in the *cis* stability is observed in half-lives between azobenzene and an azobenzene-ester, (E,Z)-ethyl 4-(phenyldiazenyl)benzoate (ABET), in DMSO at 35°C, and are consistent with previous literature findings[26]. Another distinctive trend for azobenzene-dipeptide conjguates is an increase in half-life as the number of ortho-fluorine substituents increases. The mechanism behind thermal stability of *ortho*-fluorine substituted azobenzenes is due to destabilization of the transition state and differences between the dipole moment of the *cis* isomer and the transition state. Fluorine substitutents in the *ortho* position increase the dipole moment of the *cis* isomer, and have less effect on the dipole moment of the transition state, thereby increasing the energetic potential and barrier of isomerization.[27]

III. LOGP DETERMINATION OF AZO-DYES

3.1 Solubility Studies of Azo-Dyes via UV-Vis Spectroscopy

In order to further investigate the increased *cis* stability of azobenzenes in water, the aqueous solubility *cis* and *trans* photoisomers was measured according to a method originally reported by Hartley.[5] 1 gram of azobenzene was dissolved in toluene and left in sunlight for 6 hours. The mixture of isomers were then separated and purified with flash chromatography on silica gel (32-63 μm, 60 Å pore size), solvent was removed and samples were stored at 0°C prior to use. Solubility measurements were taken by adding an excess of material to 100 mL of ddH2O. The solutions were then sonicated for 30 minutes, and filtered. Filtered solutions were used for absorbance spectroscopy, and concentrations were obtained using molar absorbtivity values from literature.[28]

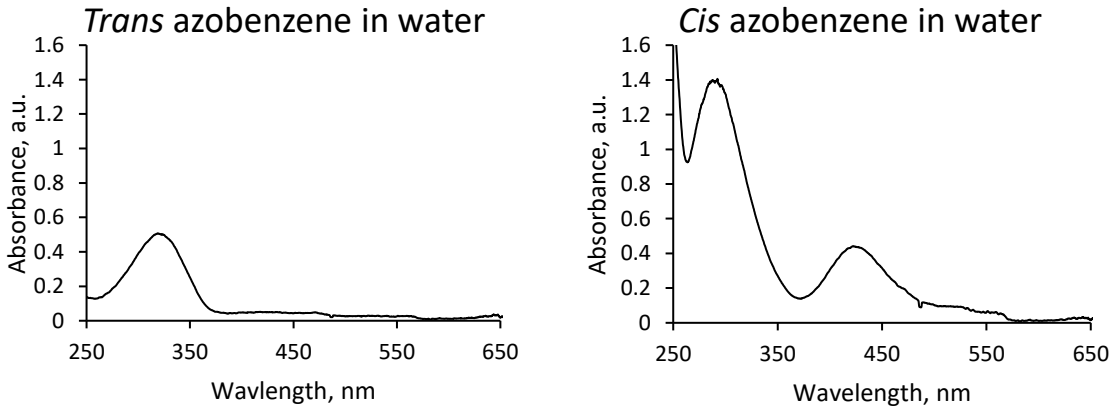


Figure 9 Ultraviolet-visible absorption spectra for saturated aqueous solutions of *trans* azobenzene (left) and *cis* azobenzene (right).

Absorption spectra of *trans* azobenzene has a distinct π-π* absorbance maximum at 325nm with a value of 0.6 a.u. The spectra of *cis* azobenzene shows two maxima, with

π-π* observed at 290nm and n-π* observed at 430nm. The literature reports molar absorbtivity values of 10000 M-1 cm-1ll for *cis* azobenzene, and 30000 M-1 cm-1 for *trans* azobenzene.[28] Using beer's law, the concentration of the samples were calculated at 0.21mM and 0.02 mM for *cis* and *trans* respectively.

3.2 LogP of Azobenzene in Water

LogP values provide information about the aqueous solubility and lipophilicity of azobenzenes. LogP values are obtained through either theoretical or experimental methods. The theoretical method involves assigning a numerical value to certain functional groups and comparing that to structures with known values. Experimental methods for determination involve placing octanol and water into a flask with the solute, shaking the container, and allowing the phases to separate into two layers. These layers are then isolated and solvent is removed. The remaining material is weighed and the logarithm of the ratio between solute in octanol and water is taken to obtain the LogP. The theoretical method does not account for the differential solubility of *cis* and *trans* photoisomers, while the experimental method is time consuming and error prone. Azobenzene is likewise sparingly soluble in water, such that the amount of water and octanol required to obtain significant recovery of material and meaningful values for LogP would be cost prohibitive. *Cis* azobenzenes are also not thermodynamically stable, and therefore the amount of time required to measure LogP through experimental means would not allow for reliable solubility values.

3.3 LogP Determination Via HPLC Retention Time Correlation

A method of LogP determination using reverse-phase HPLC developed by Abraham and coworkers[29,30] was used to measure azobenzene derivatives. The method uses a combination of experimental and computational approaches in order to correlate partitioning equilibria to HPLC retention times. The physical basis of this method relies on linear free energy relationships, which are used extensively in physical sciences to relate two similar equilibrium processes.[30]

The general form of a linear free energy relationship is described as

$$\ln K_1 = m * \ln K_2 + c \qquad \textbf{Equation 8}$$

where there is a linear relationship between the natural log of an equilibrium constant for one reaction series, K_1, to those of a second reaction series, K_2, given that either reactants or system conditions remain constant. For this application, the relationship between a solute's LogP and it's retention time, t_R will be examined.

$$\text{LogP} = \text{Log}\left(\frac{[solute]_{octanol}}{[solute]_{water}}\right) \qquad \textbf{Equation 9}$$

LogP, is the logarithm of the partition coefficient, P, for a solute between two immiscible phases, in this case octanol, and water. P is an equilibrium constant for when a solute's distribution between two phases is in equilibrium. The partitioning of a solute between two immisible phases is a mass transfer process, where a solute will move from an area of high chemical potential to low chemical potential until the system reaches chemical equilibrium.

$$\mu_i = \left(\frac{\partial \mathrm{G}}{\partial \mathrm{N}_i}\right)_{\mathrm{T},P,\mathrm{N}_{j\neq i}} \qquad \textbf{Equation 10}$$

The chemical potential, μ, of species *i*, is the partial derivative of the Gibb's free energy with respect to N number of moles of species i, under constant temperature, T, pressure, *P*, and N moles of other components. When the values of T, P, and N_j remain constant for a given process, the partial derivative, ∂G can be treated as the differential, dG.

$$\Delta \mathrm{G}_{\mathrm{octanol}\rightarrow\mathrm{water}} = \Delta \mathrm{G}^{\circ}_{\mathrm{octanol}\rightarrow\mathrm{water}} + R \cdot \mathrm{T} \cdot \ln \mathrm{Q} \qquad \textbf{Equation 11}$$

Under the previous definitions, and written in the common form, the gibb's free energy at any point for the mass transfer process, $\Delta G_{octanol\text{-}water}$, is equal to the standard-state free energy, $\Delta G^{\circ}_{octanol\text{-}water}$, plus the product of gas constant, *R*, temperature, T, and natural log of reaction quotient Q. At equilibrium, $\Delta G_{octanol\text{-}water}$ is equal to zero, and the equation is rewritten as

$$\Delta \mathrm{G}^{\circ}_{octanol\rightarrow water} = R \cdot \mathrm{T} \cdot \ln \mathrm{K}_{octanol\rightarrow water} \qquad \textbf{Equation 12}$$

$$\mathrm{K}_{octanol\rightarrow water} = \frac{\mathrm{N}_{octanol}}{\mathrm{N}_{water}} \qquad \textbf{Equation 13}$$

Where the standard free energy, $\Delta G^{\circ}_{octanol\text{-}water}$, is equal to the product of R, T, and the natural log of equilibrium constant, K, which is the ratio of N number of moles of a chemical species in octanol ($N_{octanol}$) and water (N_{water}).

The retention time, t_R, of a solute is a result of its interaction between the aqueous mobile phase and non-polar stationary phase during column chromatography,

$$k_r = \left(\frac{t\mathrm{R} - \mathrm{t}_0}{\mathrm{t}_0}\right) \qquad \textbf{Equation 14}$$

Where retention factor, k_r, is the relationship between solute retained to the stationary phase expressed as t_R, and an unretained solute that remains in the mobile phase, given as the void volume t_0. Similar to the mass transfer process described in **Equation 9**, a solute will move from an area of high chemical potential to an area of low chemical potential until equilibrium is met.

$$\frac{k_{adsorption}}{k_{desorption}} = \mathrm{K}_{stationary \rightarrow mobile} \qquad \textbf{Equation 15}$$

When the exchange of mass between mobile and stationary phase are in equilibrium, the ratio of the rate of adsorption, $k_{adsorption}$, to the desorption rate $k_{desorption}$ is equal to the equilibrium constant written as $K_{stationary->mobile.}$ Since the concentration of solute is below column capacity, the adsorption isotherm is linear, and can be related to retention factor, k_r,

$$\mathrm{k}_r = \phi \cdot \mathrm{K}_{stationary \rightarrow mobile} \qquad \textbf{Equation 16}$$

Where k_r is equal to the product of the phase ratio, ϕ, and the equilibrium constant $K_{stationary->mobile}$. Similar to **Equation 11,** the standard-state free energy of adsorption, $\Delta G°_{stationary \rightarrow mobile}$, can be expressed in the identical form of

$$\Delta \mathrm{G}°_{\mathrm{stationary} \rightarrow \mathrm{mobile}} = R \cdot \mathrm{T} \cdot \ln \mathrm{K}_{stationary \rightarrow mobile} \qquad \textbf{Equation 17}$$

Under isothermal conditions the standard state free energies remain constant and are equal to the natural log of their equilibrium constants. Since the equilibrium constants, $K_{octanol \rightarrow water}$, and $K_{stationary \rightarrow mobile}$ have been shown to be linear functions of previously defined experimental variables P, and t_R, a linear free energy relationship is defined in the form of **Equation 7** where

$$\text{LogP} = x * t_R + c \qquad \textbf{Equation 18}$$

Where LogP is equal to the retention time, t_R, times coefficient x, plus intercept, c. t_R is treated as a solute descriptor, while the coefficient x describes the contribution of t_R to the LogP value.

3.4 Experimental Methods for LogP Determination

A training set of reference compounds with known LogP values was generated using retention times experimentally derived with reverse-phase HPLC. All samples were prepared by dissolving a small amount of material in chloroform, adding several drops to a clean glass vial, and allowing the chloroform to evaporate leaving a film. Next a solution of 10% methanol and water, HPLC grade, was added to the glass vial. The mixture was then sonicated for 30 minutes, and filtered through 25µM GFP syringe filters prior to HPLC injection. Experiments were performed on an Agilent Prostar® high performance liquid chromatograph, with a Rheodyne® manual sample injector and 100µL sample injection loop and 30uL sample volume. A waters Delta-Pak® c18 300Ä column was used with 3.9mm diameter and 300mm length. The injector and lines were flushed with 100% methanol and 10% methanol/water mixtures prior to each run. Chromatography used a gradient elution method starting with a 10% methanol, 90%

water mixture and increasing to 100% methanol over a period of 50 minutes. After each run, the column was allowed 10 minutes to equilibrate to the initial 10% methanol/water mixture before injecting more samples. Reference compounds used consisted of aromatic moeities containing major functional groups i.e. phenols, alcohols, amides, amines, ketones, fluorinated aromatics. Retention times were taken with respect to peak maxima on the chromatogram, and plotted against literature LogP values. Literature LogP values were obtained from NCBI PubChem and NIST data tables. Experimentally derived LogP values were used except in instances where there were none available.

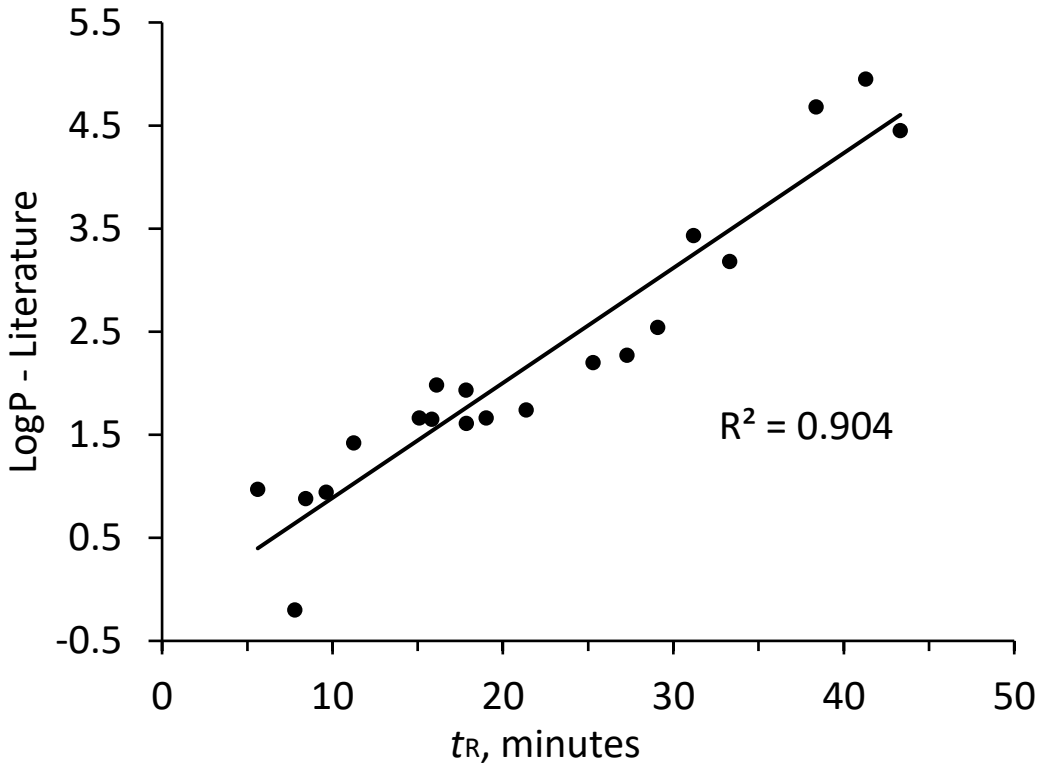


Figure 10. Plot and linear regression of HPLC retention times, and literature LogP values.

A plot of literature LogP values and retention times (**Figure 10**), and its regression provided a basis linear function for predicting LogP;

$$\mathrm{LogP_{calc}} = 0.112(\pm 0.009) * t_R - 0.228(\ \pm 0.209) \qquad \textbf{Equation 19}$$

n = 20 r = 0.951 rms = 0.405

Where $\mathrm{LogP_{calc}}$ is the calculated partition coefficient for a solute in octanol and water, and tR is the retention time of a compound in minutes. The correlation between retention time and $\mathrm{LogP_{calc}}$ is significant (**Figure 10**), however there is a relatively small available sample size of n=20 and standard error of the intercept of the line is considerably large. According to the root-mean-square error (rms), the estimation of $\mathrm{LogP_{calc}}$ is within 0.405 log units (**Equation 19**). From this basis set, there is some degree of uncertainty of the estimated values, likely the result of variables that have not been accounted for. Under the premise of solvation theory, the quantity of residuals observed would be expected, due to nearly all of the training set molecules containing varying degrees of hydrogen bond donors and acceptors. The ability of a solute to form intramolecular hydrogen bonds lowers the chemical potential of the mass transfer process described in **Equation 9** and $\mathrm{LogP_{calc}}$ values will be lower than their actual value.[30]

The plot of residual values and the value of the x-variable, $\mathrm{LogP_{calc}}$, from **Figure 11** is non-random and U-shaped, further suggesting that this current model is non-linear and requires further optimization.

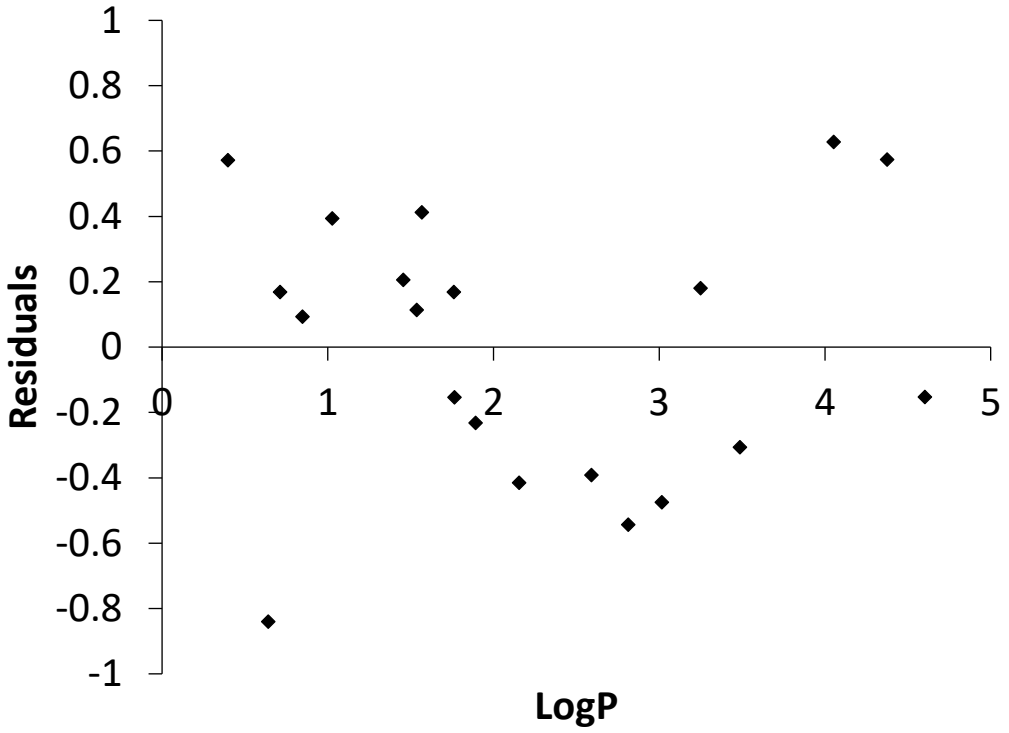


Figure 11. Plot of residuals for $LogP_{calc}$ from **Equation 19**.

Looking at the training set there is one outlier observed, 4'-amino-benzylalcohol the difference between the reported LogP and $LogP_{calc}$ is 0.84 log units. This data point was removed from the sample population, and was improved with r = 0.955, and an estimation rms = 0.360.

Previous work has shown that hydrogen bond donors and acceptors have significant contributions to the solvation properties of a chemical species.[32,33] When hydrogen bond donor and acceptor count terms are applied to this model, and a plot of $LogP_{calc}$ values derived from t_R and LogP literature values are made, a stronger correlation is observed in **Figure 12**.

Table 4. Average retention times, reference LogP values, and values calculated from **Equation 19** for compound training set.

Name	Avg Tr	LogP -ref	$LogP_{calc}$ eq19
fmoc-Leu-OH	**41.28**	**4.95**	**4.38**
2-phenyl-indole	**38.38**	**4.68**	**4.05**
1-methyl-2-phenylindole	**43.32**	**4.45**	**4.60**
bisphenol A	**31.18**	**3.43**	**3.25**
2-phenyl-acetophenone	**33.31**	**3.18**	**3.49**
isobutyrlphenone	**29.08**	**2.54**	**3.01**
4'-methoxy-propiophenone	**27.27**	**2.27**	**2.81**
propiophenone	**25.28**	**2.20**	**2.59**
2,6-difluorophenol	**16.10**	**1.98**	**1.57**
3-fluorophenol	**17.84**	**1.93**	**1.76**
p-methoxyacetophenone	**21.37**	**1.74**	**2.16**
acetanilide	**15.09**	**1.66**	**1.45**
acetophenone	**19.01**	**1.66**	**1.89**
2,6-difluoroaniline	**15.82**	**1.65**	**1.54**
fluoroacetanilide	**17.86**	**1.61**	**1.76**
4-HO-acetophenone	**11.25**	**1.42**	**1.03**
4'-amino-acetanilide	**5.61**	**0.97**	**0.40**
aniline	**9.64**	**0.94**	**0.85**
catechol	**8.43**	**0.88**	**0.71**
4'-amino-benzylalcohol	**7.79**	**-0.20**	**0.64**

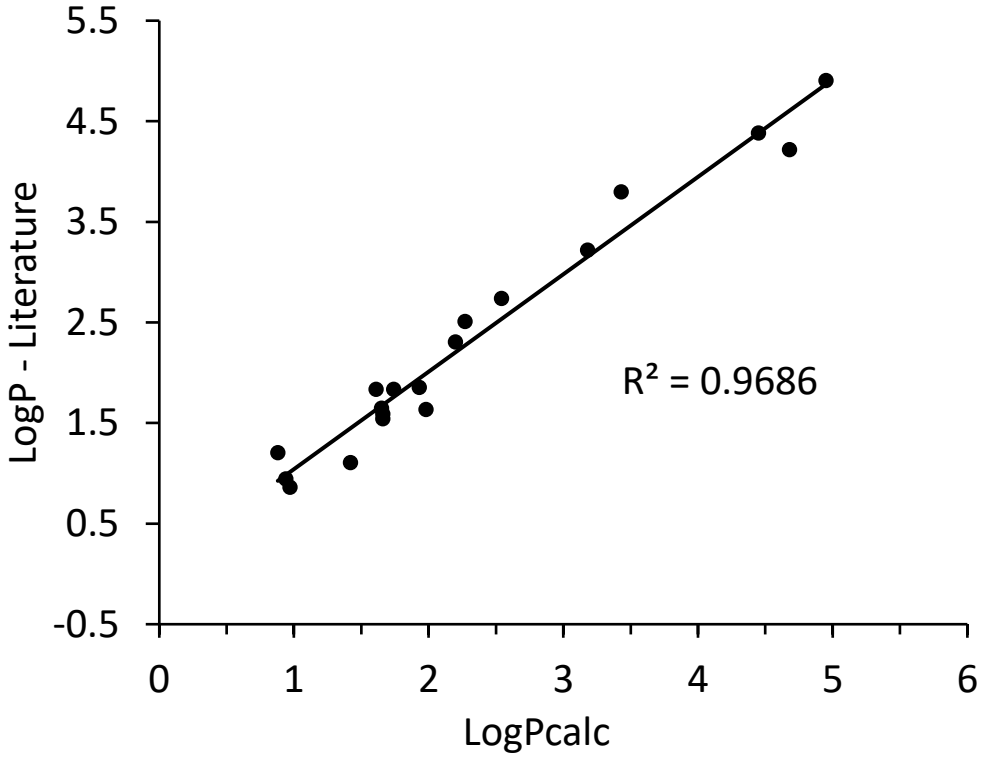


Figure 12. Plot and linear regression of $LogP_{calc}$ from **Equation 20,** and literature LogP values.

The linear function derived from the hydrogen bond donor/acceptor contribution is written as

$$LogP_{calc} = 0.113(\pm 0.005) * t_R + 0.399(\pm 0.080) * D + 0.023(\pm 0.075) * A - 0.552(\pm 0.187)$$

n = 19 r = 0.984 rms = 0.216

Equation 20

Where constants *t*R is the retention time of solute, D is the hydrogen bond donor count, and A is the hydrogen bond acceptor count. With the addition of descriptors for hydrogen bond donors and acceptors, the correlation is improved indicated by the LogPcalc estimation falling within 0.22 log units. The hydrogen bond acceptor term, A, is clearly an outlier as indicated by it's standard error (**Equation 20**). Here it can be

observed that the most significant coefficient contribution comes from the hydrogen bond donor term. The residuals in (**Figure 13**) are more random, also supporting improved correlation.

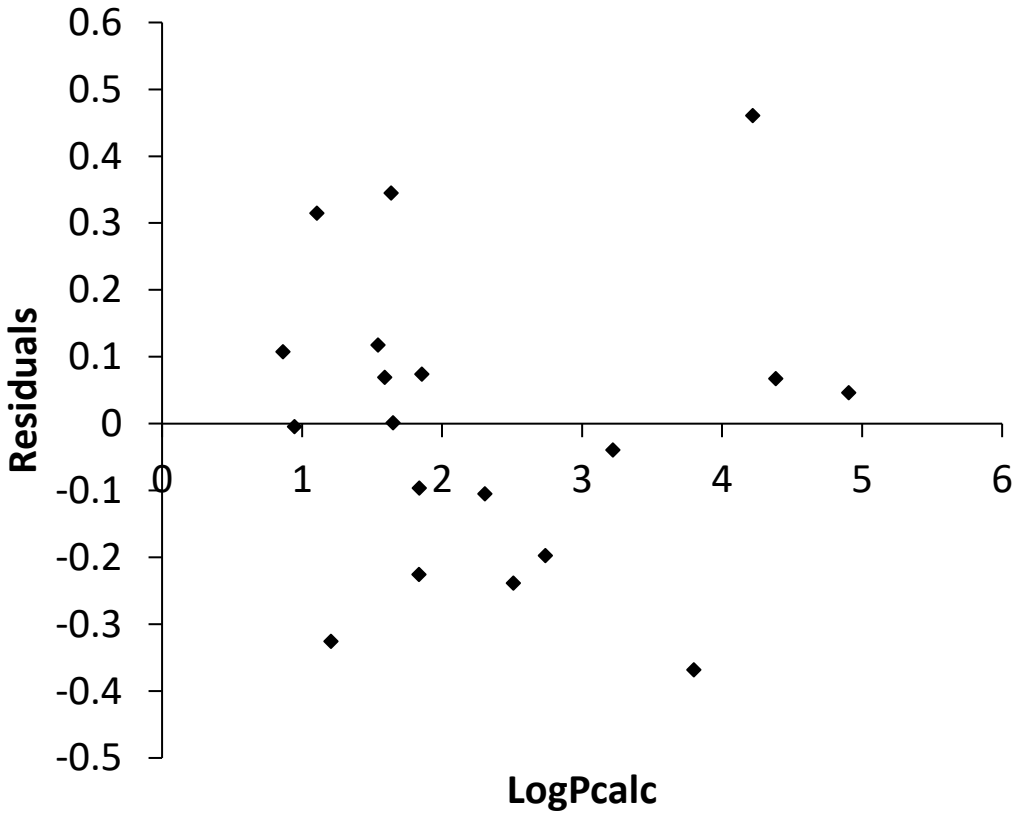


Figure 13. Plot of residuals of $LogP_{calc}$ from **Equation 20**

A comparison of standard errors in the function indicates that the hydrogen bond acceptor term, A, is likely insignificant.

$$LogP_{calc} = 0.100(\pm 0.004) * t_R + 0.293 \pm (0.043) * D - 0.327 \pm (0.112)$$

n = 19 r = 0.976 rms = 0.267

Equation 21

An equation was generated correlating only two terms, retention time, t_R and hydrogen bond donor count, D. Exclusion of the hydrogen bond acceptor term produced a linear regression with an Rsqr of 0.952. The bias of the linear regression of **Equation 21** was tested by randomly removing a data point from the sample set and looking at regression results. Testing was performed in triplicate, and the variance of the regression coefficients are shown in **Table 5**.

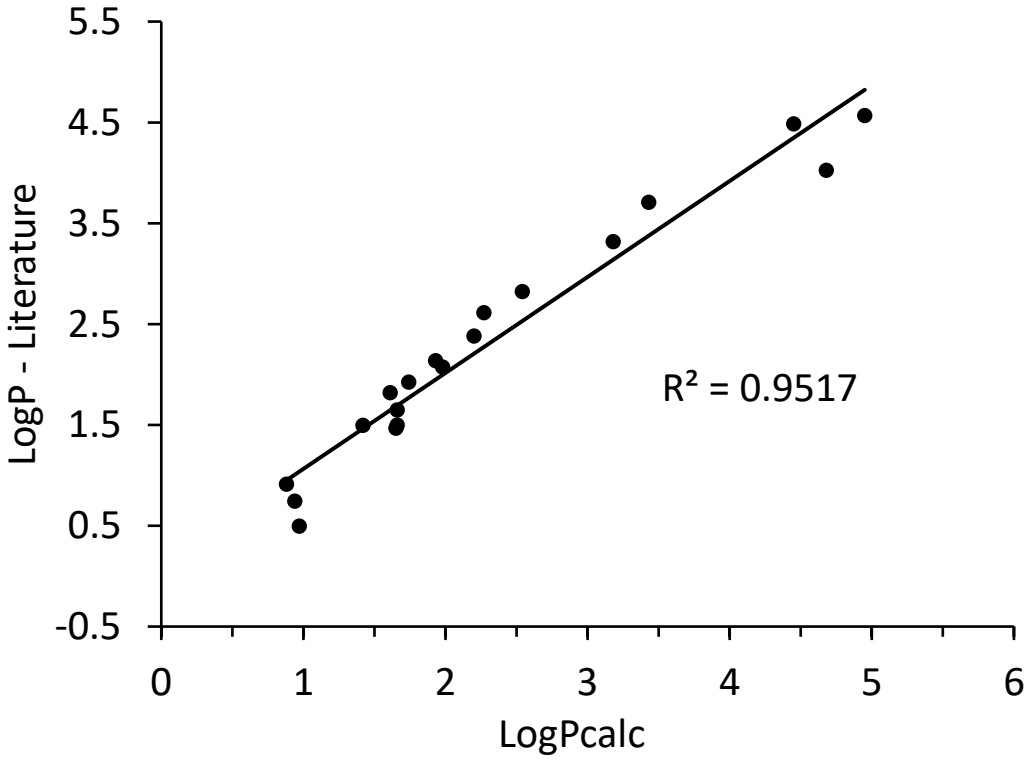


Figure 14. Plot and linear regression of LogP literature values and $LogP_{calc}$ values obtained from **Equation 21**.

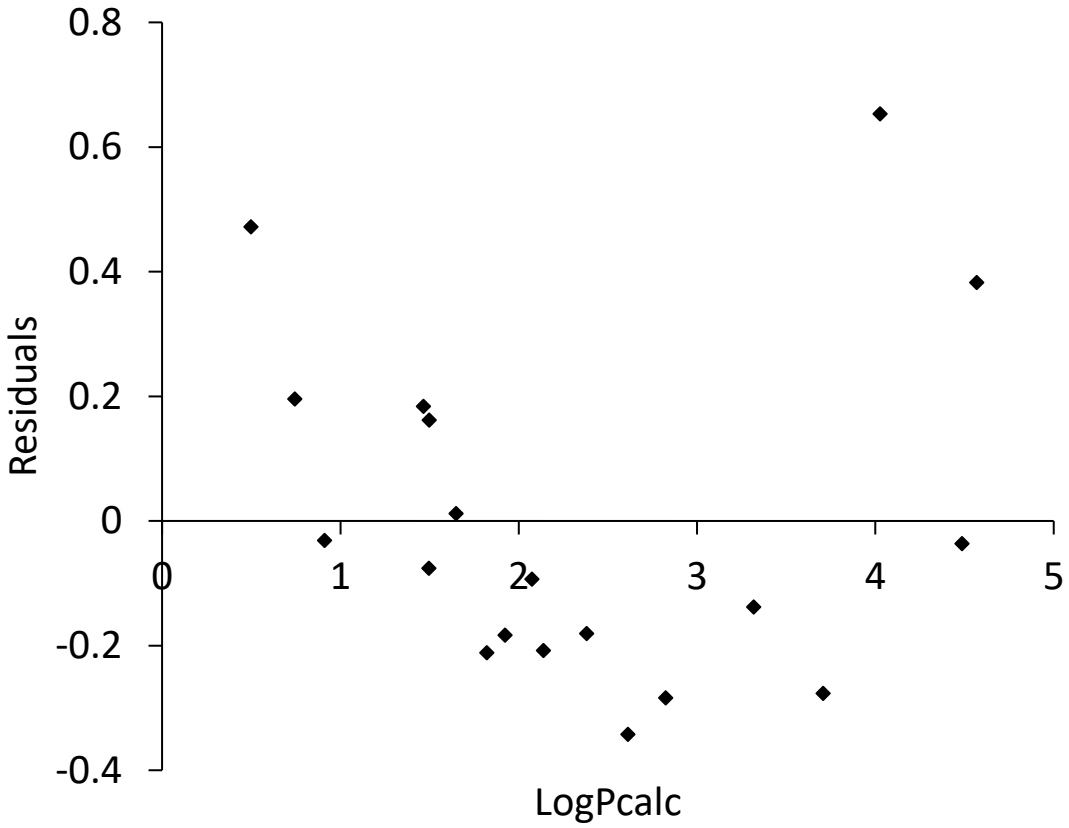


Figure 15. Plot of residuals of $LogP_{calc}$ from **Equation 21**

3.5 Results for LogP Determination of Azo Dyes

$LogP_{calc}$ values generated from **Equation 20** and **Equation 21** were compared to LogP values obtained from literature (**Table 5**). The calculated variance between all $LogP_{calc}$ and LogP literature values were 0.011 for **Equation 20** and 0.022 for **Equation 21**. For both equations, the largest outliers were observed towards compounds with LogP values greater than 4. For **Equation 20**, there is a discrepancy between literature and calculated methods for 2-phenyl-indole, with the literature LogP and $LogP_{calc}$ value differing by 0.46 log units. Similarly, the calculated LogP for 1-methyl-2-phenylindole

from **Equation 21** differs from literature values by 0.61. As expected, the regression results has the least variance with compounds in the middle of the regression.

Table 5. Comparison of LogP literature values, $LogP_{calc}$ values from **Equation 20** (eq20), and **Equation 21** (eq21).

Name	tR (min)	LogP literature	LogP eq20	ΔLogP eq20	LogP eq21	ΔLogP eq21
1-methyl-2-phenylindole	43.32 ± 0.07	4.45	4.38	0.07	3.79	0.61
fmoc-Leu-OH*	41.28 ± 0.14	4.95	4.90	0.05	4.02	0.08
2-phenyl-indole	38.38 ± 0.14	4.68	4.22	0.46	3.82	0.02
bisphenol A	31.18 ± 0.78	3.43	3.80	0.37	3.44	0.01
2-phenyl-acetophenone	33.31 ± 0.14	3.18	3.22	0.04	3.00	0.20
isobutyrlphenone	29.08 ± 1.4	2.54	2.74	0.20	2.69	0.15
4'-methoxy-propiophenone	27.27 ± 1.32	2.27	2.51	0.24	2.50	0.23
propiophenone	25.28 ± 0.93	2.20	2.31	0.11	2.27	0.07
2,6-difluorophenol	16.10 ± 0.25	1.98	1.64	0.35	1.56	0.42
3-fluorophenol	17.84 ± 1.18	1.93	1.86	0.07	1.67	0.26
p-methoxyacetophenone	21.37 ± 0.88	1.74	1.84	0.10	1.88	0.14
acetophenone	19.01 ± 1.48	1.66	1.59	0.07	1.68	0.02
acetanilide	15.09 ± 1.27	1.66	1.54	0.12	1.39	0.27
2,6-difluoroaniline	15.82 ± 0.17	1.65	1.65	-	1.86	0.21
fluoroacetanilide	17.86 ± 0.58	1.61	1.84	0.23	1.80	0.19
4-hydroxy-acetophenone	11.25 ± 1.02	1.42	1.11	0.31	1.50	0.08
4'-aminoacetanilide	5.61 ± 0.29	0.97	0.86	0.11	0.86	0.11
aniline	9.64 ± 0.85	0.94	0.94	-	0.99	0.05
catechol	8.43 ± 0.42	0.88	1.21	0.33	1.13	0.25

A comparison of coefficient values for **Equation 19**, **20** and **21** (**Table 6**) indicates that the hydrogen bond acceptor term is not only the least significant contribution, but also has a negligible contribution to LogP values. The hydrogen bond donor term however, has a significant contribution as a solvation parameter, and therefore **Equation 21** was chosen to be used in LogP determinations for Azo dyes.

Table 6. Comparison of coefficient values and their standard errors from Equations 19, 20, and 21.

	t	d	a	b
Equation 19	0.114 ± 0.083	-	-	-0.336 ± 0.212
Equation 20	0.101 ± 0.005	0.286 ± 0.044	0.038 ± 0.038	-0.393 ± 0.154
Equation 21	0.100 ± 0.004	0.293 ± 0.043	-	-0.327 ± 0.122

The $LogP_{calc}$ values of a library of azobenzenes was then obtained using **Equation 21, Figure 14** derived from the training set previously mentioned (**Table 7**). Fluoro-substituted methoxy azobenzenes (**Scheme 2**), and azobenzene bioconjugates (**Scheme 3**) were prepared by dissolving in a solution of 10% methanol/water, sonicating, and irradiating with a Hg/Xe arc lamp, unfiltered for 30 minutes. Photoirradiation by this method was sufficient to produce a mixture of *cis* and *trans* photoisomers, which had differential retention times and detection by a UV absorbance detector. Experimental LogP values were obtained for azobenzene and 4'-methoxyazobenzene from NCBI Pubchem database. These values obtained did not provide LogP values for *cis* azobenzene. The mean difference in LogP values for the entire set of azobenzenes tested was 1.01 log units. All of the azobenzenes tested exhibited a higher $LogP_{calc}$ value for *trans* compared to *cis* except for CF3AB (trifluoromethyl-azobenzene). All of the methoxy-azo dyes showed similar retention time behavior for the *trans* isomer except for Fa1ABOMe, which eluted considerably faster. The lower $LogP_{calc}$ value for Fa1ABOMe

may be due to a greater dipole moment from the electron-donating methoxy group and flourine group being located on the same ring system. The peptide-based azobenzene conjugates had similar differences between *cis* and *trans* $LogP_{calc}$ values as seen with methoxy azo-dyes, but the *trans* isomers for the dipeptide based materials had a $LogP_{calc}$ value close to 1 log unit less. These peptide based materials were observed to be slightly more soluble, and furthermore pH was observed to have a significant effect on peptide conjugate solubilities.

Table 7. Comparison of $LogP_{calc}$ values for azo-dyes generated using **Equation 20**.

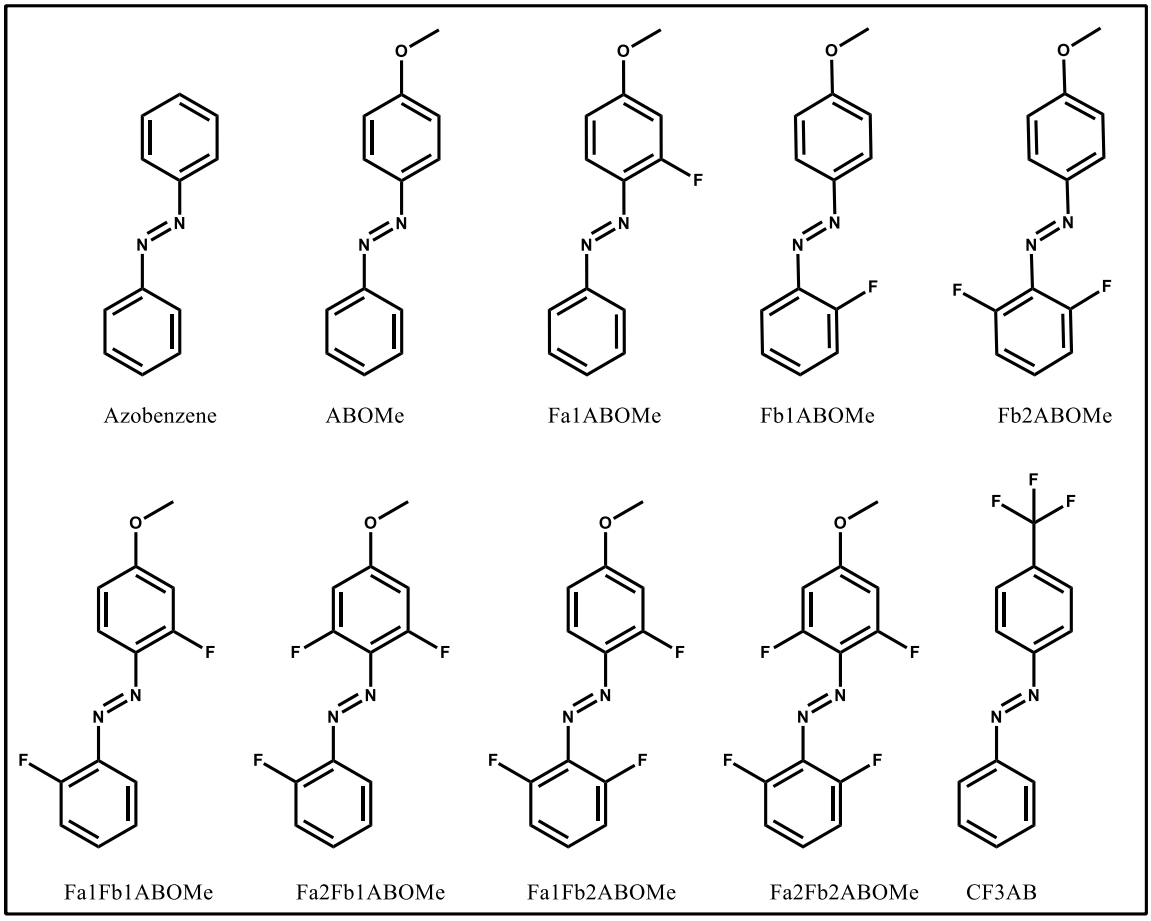
Compound	LogP, *trans*	LogP, *cis*	ΔLogP
Azobenzene	4.19	2.80	1.39
ABOMe	4.22	2.9	1.32
Fa1ABOMe	3.68	2.31	1.36
Fb1ABOMe	4.17	2.86	1.31
Fb2ABOMe	3.97	3.02	0.95
Fa1Fb1ABOMe	4.24	3.05	1.19
Fa1Fb2ABOMe	3.96	3.11	0.84
Fa2Fb1ABOMe	4.05	3.26	0.80
Fa2Fb2ABOMe	3.82	3.27	0.56
CF3AB	4.53	4.60	-0.07
ABDP	3.03	1.92	1.11
ABPE	3.97	2.96	1.01
RABDP	2.67	1.79	0.87
1FABDP	3.18	2.21	0.97
2FABDP	3.37	2.38	0.99
ABNONA	0.53	-0.18	0.71

Comparisons of dipole moments calculated using DFT with basis set b3lyp at 6-31g d,+ (**Table 8**) shows that for CF3AB, the dipole moments are switched compared to typical azo-dyes, with the *trans* isomer having a greater dipole moment compared to the *cis.*

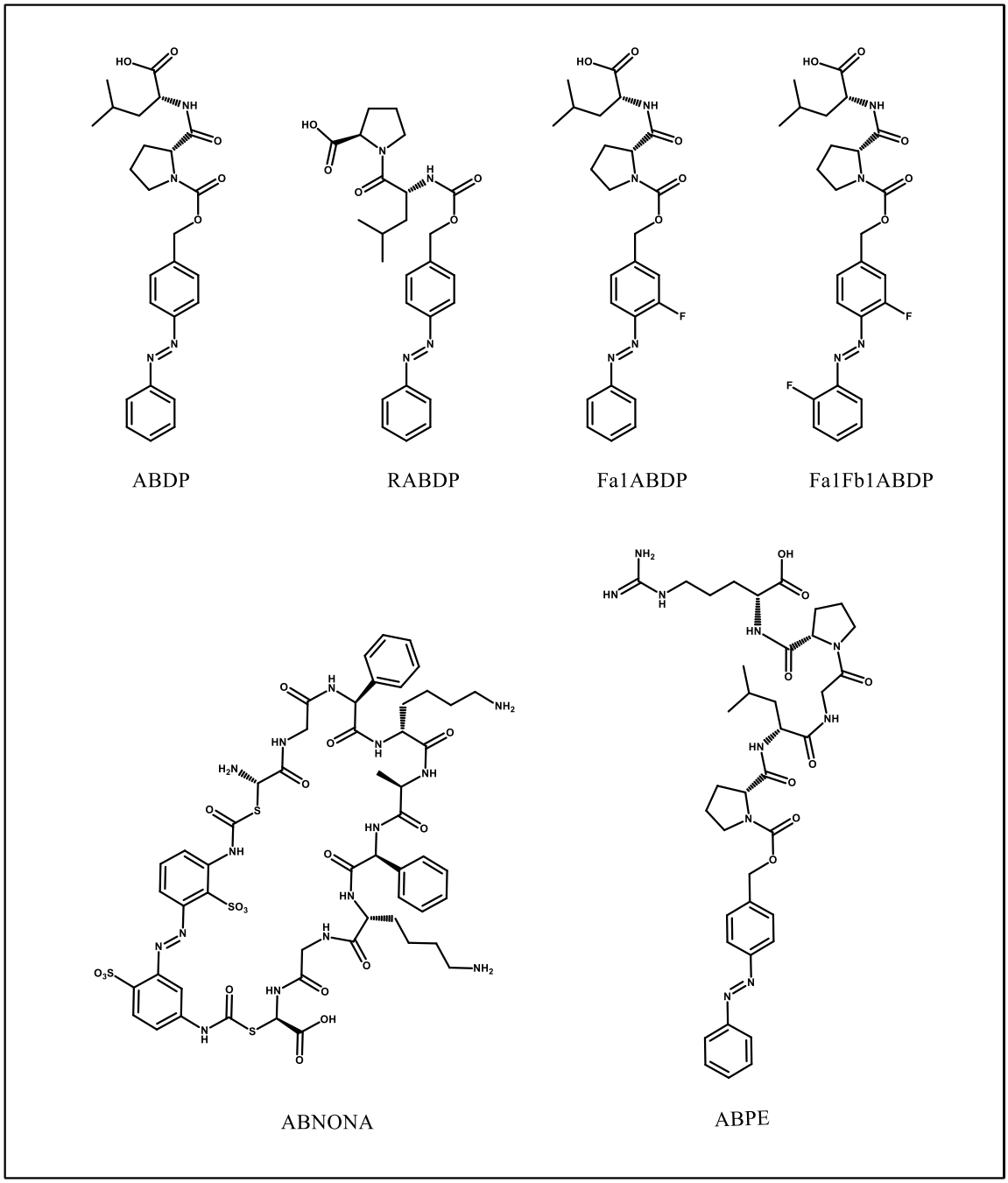
Table 8. Comparison of DFT calculated dipole moments (μ), and $LogP_{calc}$ values from **Equation 20**.

Compound	μ, *trans*	μ, *cis*	Δμ	LogP, *trans*	LogP, *cis*	ΔLogP
Fb2ABOMe	2.00	6.80	4.80	3.83	2.84	0.99
Fa2Fb1ABOMe	2.90	7.60	4.70	3.95	3.01	0.94
Azobenzene	0.00	4.60	4.60	3.92	2.63	1.29
ABDP	5.80	10.10	4.30	3.55	2.47	1.07
CF3AB	4.80	3.70	-1.10	4.24	4.30	-0.07

The dipole moments for the peptide-based ABDP was calculated by taking an average of optimized energy minima for several conformations. The calculated dipole moment for ABDP is substantially higher for the *trans*, owing to the conformational flexibility and extended size of the molecule, and may be a limitation of computational methods. More importantly, the differences in solubility and $LogP_{calc}$ between the photoisomers of these molecules, is attributed to the structural and electronic differences between *cis* and *trans* which mediates a change in the polarity of the compound.



Scheme 2. Molecular structures of azobenzene, and fluoro-substituted methoxy azobenzenes used for $LogP_{calc}$ determinations.



Scheme 3. Molecular structures of azobenzene bioconjugates used for $LogP_{calc}$ determination.

IV. CONCLUSIONS

Azobenzene-dipeptide derivatives were designed with *ortho* fluorine substituents and the dipeptide sequence reversed in order to examine stability of the *cis* photoisomer. These modifications were based off of previous findings,[6,12] which observed increased *cis* stability in *ortho* fluorine substituted azobenzenes, and in an azobenzene-peptide conjugate. ^{1}H NMR of the reverse dipeptide, RABDP, revealed similar *cis*:*trans* ratios in D_2O as what was observed in the ABDP molecule, and indicated that sequence order has a negligible effect on photostationary *cis*:*trans* ratios. The behavior of *ortho* fluorine derivatives were not able to be analyzed in water by ^{1}H NMR due to increased S/N and marginal solubility. The rates of the thermal *cis* - *trans* isomerization in methanol and water were not significantly different between ABDP and RABDP, and supports stabilization through an electron donor effect rather than a steric contribution suggested by Haworth[6].

Differential solubility between *cis* and *trans* photoisomers was assessed by determining the LogP of azobenzene-peptide bioconjugates and fluorinated *para*-methoxyazobenzenes. A method using RP-HPLC retention time correlations permitted LogP determination of individual *cis* and *trans* isomers. The introduction of fluorine substituents in the *ortho* position did not substantially effect the $LogP_{calc}$ of the *trans* isomers, but significant variability was observed in the retention times and $LogP_{calc}$ of *cis* isomers that ranged between 2.22 to 3.07 log units for *para*-methoxy azo dyes and 2.47 to 3.28 for azobenzene-dipeptides. For azobenzene dipeptide conjugates, $LogP_{calc}$ of *cis* and

trans isomers increased with addition of flourine substituents due to increases in hydrophobicity. The average change in $LogP_{calc}$ between *cis* and *trans* isomers was 1.13 log units. According to computational approaches,[34] this implies on average there is a 14-fold change in aqueous solubility between photoisomers for azo-compounds tested. Significant light-induced changes in solubility can be applied to direct modulation of intensive properties of systems at interfaces such as surface tension, membrane permeability,[35] and wetability. *DFT* calculations of the dipole moment for a *para*-trifluoromethylazobenzene showed a reverse trend where the *trans* isomer was more polar than the *cis* isomer. Similarly, the *trans* isomer of CF3AB eluted first in RP-HPLC studies, and the values of the retention times of *cis* and *trans* were respectively 47.29 and 45.45 min.

APPENDIX SECTION

A. Synthesis of Azobenzene-Dipeptide Conjugates

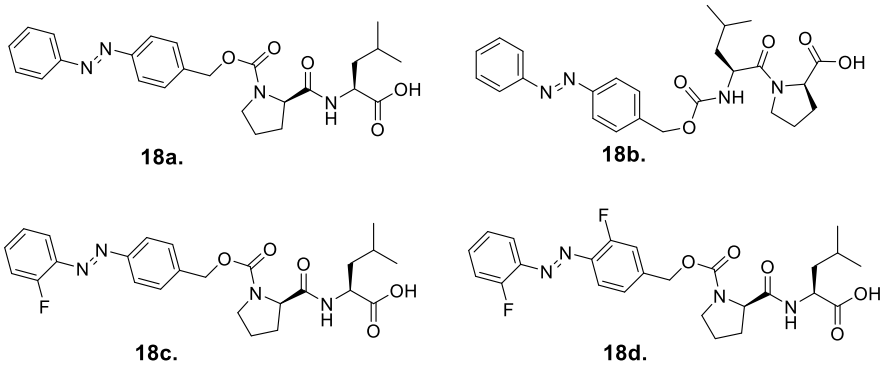


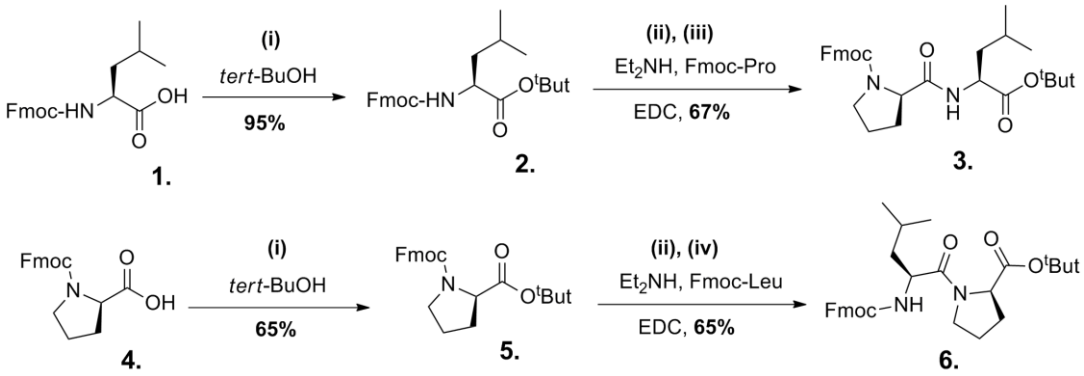
Figure A1: Structures for azobenzene-dipeptide conjugates (**18a-d**); ABDP (**18a**), RABDP (**18b**), 1FABDP (**18c**), 2FABDP(**18d**).

A1.1 Reagents, Materials and Instruments

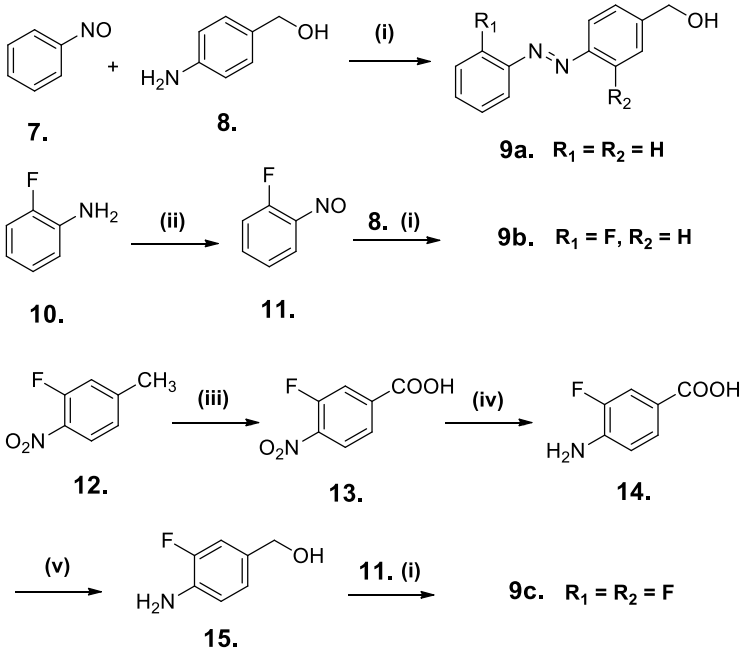
Unless otherwise noted, all reagents and solvents were purchased from commercial sources (Acros Organics and Sigma-Aldrich, USA) and used without purification. Dry solvents were obtained using a solvent purification system (Innovative Technology, Inc.). Chemical reactions were performed in oven-dried flasks open to the atmosphere or under nitrogen gas and monitored by thin layer chromatography (TLC). TLC was performed on precoated silica gel 60 F254 (250 μm) plastic-backed plates (EMD Chemicals Inc.) and visualization was accomplished with UV light (254 nm). Flash column chromatography was performed on silica gel (32-63 μm, 60 Å pore size). NMR spectra were recorded on Bruker Avance III 400 and 500 MHz spectrometer. ^{1}H and ^{13}C NMR shifts are reported relative to TMS and ^{19}F shifts are relative to C_6F_6 (-

164.9 ppm). (400.13 MHz for ^{1}H; 100.61 MHz for ^{13}C; 376.46 MHz for ^{19}F) Chemical shifts (δ) are reported in ppm relative to the TMS internal standard. Abbreviations are as follows: s (singlet), d (doublet), t (triplet), q (quartet), m (multiplet) and coupling constant (*J*) in Hz. ESI-MS (sensitivity mode) analyses were performed using Waters Synapt G2 mass instrument.

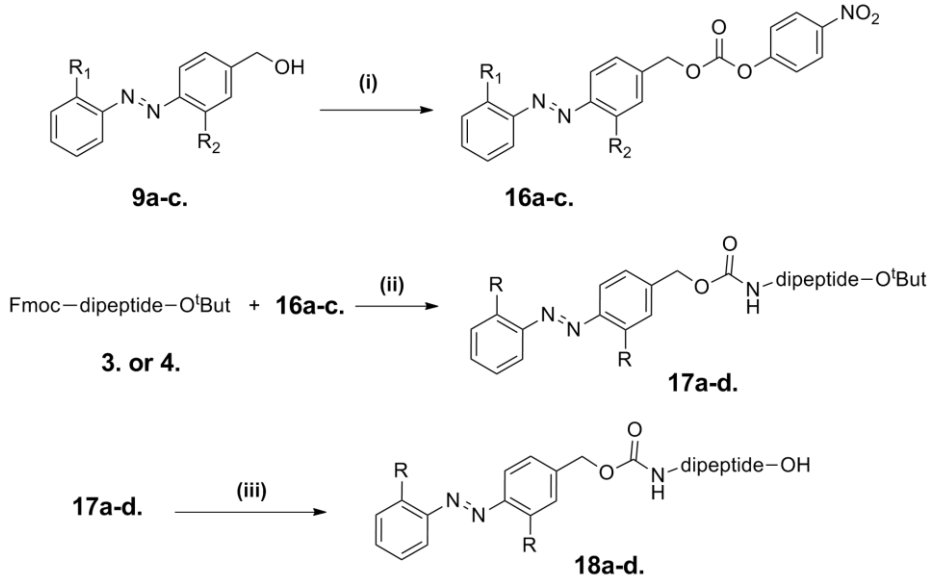
A1.2 Synthetic Schema and Methods



Scheme A1: Synthetic paths for Fmoc, and *t*-But protected dipeptides **3** and **6.** (i) **1.** Fmoc-Leu-OH, or **4.** Fmoc-Pro-OH, *t*-BuOH, DCC, DMAP and Dry DCM, 0°C to rt 16h, (ii) 20%(iPr)$_2$NH, ACN, rt, 4 h, (iii) Fmoc-Pro-OH, DCC, DCM, 0°C to rt, 4h, (iv) Fmoc-Leu-OH, DCC, DCM, 0°C to rt, 4h.



Scheme A2: Synthetic paths for (*E*)-(4-(phenyldiazenyl)phenyl)methanol **9a** and *ortho*-mono-fluoro **9b** and *ortho*-difluoro derivatives **9c.** (i) AcOH (glacial), r.t., 16-24h, (ii) Oxone-monosulfate, DCM/H_2O, r.t., 24h. (iii) $Na_2Cr_2O_7.2H_2O$, H_2O/H_2SO_4, r.t., 16h, Δ 2h, (iv) 0.1 mole% of 5% pd/C, Et_3N, 97% HCOOH slow addition Δ 2h, (v) $LiAlH_4$, Dry THF, 0°C to r.t. 24h.g

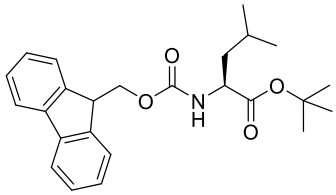


Scheme A3: General method for Pz-dipeptides (**18a-d**) synthesis: (i) *p*-Nitrophenyl-chloroformate, Et_3N, Dry THF, rt., 24h, (ii) (a) Fmoc-dipeptide-O^{t}But (**3** or **4**), 20% Et_2NH, ACN, 4h., (b) **16a-c**,, Et_3N, DCM, r.t., 24h.(iii) *p*-toluenesulfonic acid, Benzene, Δ 3h

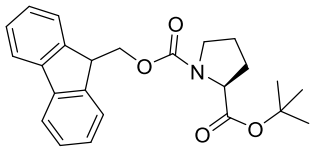
Fmoc-Leu-O^{t}But (2) and Fmoc-Pro-O^{t}But (5)

Fmoc-Leu-OH (1.06 g, 3 mmol) or Fmoc-Pro-OH (1.01g, 3 mmol) was added to a solution of *tert*-butanol (266 mg, 3.6 mmol), DMAP (73 mg, 0.6 mmol) in anhydrous DCM (50 mL) and stirred for 30 min at 0 °C. Solution of DCC (742 mg, 3.6 mmol) in DCM (10 ml) was added dropwise to above solution. Reaction mixture was warmed to room temperature and stirred for 16 h. After completion of reaction, reaction mixture was filter from celite and washed with water (10 mL twice), brine (5 mL twice) dried over

anhydrous Na_2SO_4 and concentrated under reduced pressure. The product **2** was purified by silica gel flash chromatography (EtOAc/Hexane mixtures) to provide a colorless oil in 95% yield of **2**. and an oil in 65% yield of **5.** Product **2** was characterized through ^1H NMR in ($CDCl_3$) and observed similar spectra as previously reported.[36] R_f = 0.32 (20% EtOAc/Hexane) The product **5** was also characterized through ^1H NMR in ($CDCl_3$) and observed similar spectra as previously reported and the ^1H NMR in ($CDCl_3$) of product **5** was matched with reported earlier.[37] R_f = 0.45 (20% EtOAc/Hexane)



(S)-tert-butyl 2-((((9H-fluoren-9-yl)methoxy)carbonyl)amino)-4-methylpentanoate (Fmoc-Leu-O^{t}But) **2**. ^1H NMR ($CDCl_3$): δ 7.77 (dd *J* = 0.68, 7.5 Hz, 2H, Ar-H), 7.62 (d *J* = 7.5 Hz, 2H, Ar-H), 7.40 (t *J* = 7.9 Hz, 2H, Ar-H), 7.31 (dt *J* = 1.0, 8.4 Hz, 2H, Ar-H), 5.23 (bd *J* = 8.5 Hz, 1H), 4.41 (d *J* = 1.0, 8.4 Hz, 2H, Ar-H), 4.28 (m, 1H), 4.23 (t, *J* = 7.0 Hz, 1H), 1.75-1.58 (m, 2H), 1.48 (s, 10H), 0.97 (d, *J* = 6.5 Hz, 2H), ESI-MS: m/z: calc. $C_{25}H_{32}NO_4$ [M+H$^+$] 410.2331, found 410.2540.



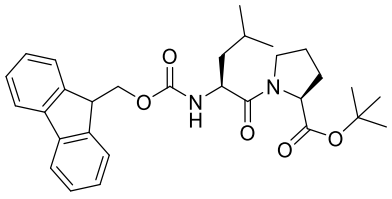
(S)-1-((9H-fluoren-9-yl)methyl 2-tert-butyl pyrrolidine-1,2-dicarboxylate (Fmoc-Pro-O^{t}But) **5**. ^1H NMR ($CDCl_3$): δ 7.77 (dd *J* = 0.68, 7.5 Hz, 2H, Ar-H), 7.69-7.58 (m, 2H, Ar-H), 7.39 (t *J* = 7.9 Hz, 2H, Ar-H), 7.31 (dt *J* = 1.0, 8.4 Hz, 2H, Ar-H), 4.46-4.44 (m, 1H), 4.35-4.17 (m, 3H), 3.68-3.61 (m, 1H), 3.57-3.51 (m, 1H), 2.28-2.16 (m, 1H), 2.08-2.04 (m, 3H), 1.47-1.44 (ds, 9H), ESI-MS: m/z: calc. $C_{25}H_{27}NO_4$ [M+H$^+$] 394.2116, found 394.2018

Fmoc-Pro-Leu-O^{t}But (3) and Fmoc-Leu-Pro-O^{t}But (6)

Deprotection of Fmoc- group: Fmoc-amino acid-ester **2** or **5** (2.9 mmol) was dissolved into a 20% (v/v) of Diethyl amine (Et_2NH) solution in CH_3CN (7 mL). The reaction

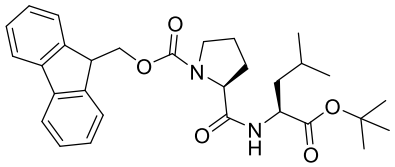
mixture was stirred at for 30 min at rt. The solvent was evaporated under reduced pressure and the crude product was used without further purification.

Coupling of amino acid: To a solution of Fmoc-aminoacid-OH (Fmoc-Pro-OH or Fmoc-Leu-OH; 3.3 mmol) in DCM (5 mL), HOBt (3.3 mmol), Et_3N (3.3 mmol) and unpurified H_2N-Leu-O^{t}But or H_2N-Pro-O^{t}But (3 mmol) solution in DCM (5 mL) were added and stirred for 10 min at 0 °C. DCC (680 mg; 3.3 mmol) solution in DCM (10 mL) was added dropwise to the amino acid mixture. The reaction mixture was warmed to room temperature and stirred for 4 hours. Completion of reaction was monitored by TLC. After completion, the mixture was filtered through a pad of cellite and the filtrate washed with 1 M HCl (5 mL), saturated solution of $NaHCO_3$ (5 mL), water (5 mL) and brine (5 mL). The organic layers were dried with anhydrous $MgSO_4$ and concentrated *in vacuo*. The product **3** was isolated by flash chromatography (25% EtOAc/hexane) to provide 67% yield as an amorphous white solid. And 20% EtOAc/Hexane was used to provide compound **6** in 65% yield as a white solid. The product **3** was characterized through ^{1}H NMR in ($CDCl_3$) and observed similar spectra as previously reported.[38] R_f = 0.38 (30% EtOAc/Hexane). The product **5** was also characterized through ^{1}H NMR in ($CDCl_3$) and observed similar spectra as previously reported and the ^{1}H NMR in ($CDCl_3$) of product **5** was matched with reported earlier. R_f = 0.39 (30% EtOAc/Hexane).



(S)-tert-butyl 1-((S)-2-((((9H-fluoren-9-yl)methoxy)carbonyl)amino)-4-methylpentanoyl)pyrrolid ine-2-carboxylate (Fmoc-Leu-Pro-O^{t}But) **3**. ^{1}H NMR ($CDCl_3$): δ 7.77 (dd *J* = 0.68, 7.5 Hz, 2H, Ar-H), 7.61 (d *J* = 7.8 Hz, 2H, Ar-H), 7.40 (t *J* = 7.9 Hz, 2H, Ar-H), 7.31 (m, 2H, Ar-H), 5.23 (bd *J* = 10 Hz, 1H, urethane NH), 4.69-4.56 (m,1H, Pro, α-CH), 4.4 (1H, Leu, α-CH), 4.35 (d, *J* = 7.0 Hz, 2H, Fmoc CH_2), 4.21 (t *J*

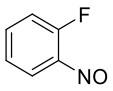
= 7,5 Hz, 1H, Fmoc CH), 3.65 (m, 1H, Pro, CH_2), 3.39 (m, 1H, Pro, CH_2), 2.16 (m, 1H, Pro CH_2), 2.04-1.95 (m, 2H, Pro CH_2), 1.82-1.68 (m, 1H, Leu γ-CH), 1.2 (m, 2H, Leu β-CH_2), 1.4 (s, 9H, Bu^t), 0.98 – 0.95 (dd, *J* = 7.5 Hz, 6H, 2 x Leu CH_3), ESI-MS: m/z: calc. $C_{30}H_{38}N_2NaO_5$ [M+Na^+] 529.2678, found 529.2780



(S)-(9H-fluoren-9-yl)methyl 2-(((S)-1-(tert-butoxy)-4-methyl-1-oxopentan-2-yl)Carbamoyl) pyrrolidine-1-carboxylate (Fmoc-Pro-Leu-O^{t}But) **6**. ^{1}H NMR ($CDCl_3$): δ 7.76 (d *J* = 7.3 Hz, 2H, Ar-H), 7.58 (m, 2H, Ar-H), 7.39 (t *J* = 7.4 Hz, 2H, Ar-H), 7.31 (dt *J* = 1.0, 8.4 Hz, 2H, Ar-H), 4.46-4.24 (m, 5H), 3.56-3.47 (m, 2H), 2.27 (m, 1H), 2.0 (m, 3H), 1.61 (m, 3H), 1.45 (d *J* = 11.1 Hz, 9H, Ar-H), 0.85-084 (ds, 6H), ESI-MS: m/z: calc. $C_{30}H_{38}N_2NaO_5$ [M+Na^+] 529.2678, found 529.0935

1-fluoro-2-nitrosobenezene (11)

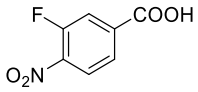
An aqueous solution of Oxone® (33.0 g, 53.7 mmol) in water (240 mL) is added to a solution of 1-fluoroaniline (2.7 mL, 27.9 mmol) in DCM (60 mL). The reaction mixture was stirred at rt. under N_2 gas with the absence of light for 24 h. After completion of reaction, the organic layer was separated and aqueous layer was extracted with DCM (20 mL) three times. The combined organic layers were washed with 5% of HCL (20 mL), saturated solution of $NaHCO_3$ (20 mL) brine, water (20 mL) and brine (10 mL), then dried over anhydrous Na_2SO_4 and concentrated under reduced pressure. The product was isolated silica gel flash chromatography (10% DCM/Hexane) and with MeOH/H_2O to provide pale yellow solid crystals 1.39 g (46% yield). ^{1}NMR in ($CDCl_3$) and observed similar spectra as previously reported.[39]



1-fluoro-2-notrosobenzene **11**. ^{1}H NMR ($CDCl_3$): δ 7.74-7.69 (m, 1H, Ar-H(5)), 7.53-7.48 (m, 1H, Ar-H (3)), 7.14 (dt J = 7.2, 1.8 Hz, 1H, Ar-H(4)), 6.51 (dt J = 7.2, 1.8 Hz, 1H, Ar-H(6)), ^{13}C NMR: δ -166.0, 154.9, 138.1, 123.7, 118.8, 109.9. ^{19}F NMR: δ -129.53.

3-fluoro-4-nitrobenzoic acid (13).

Compound **13** was synthesized using a method for oxidation of *p*-nitrotoluene.[40] To a solution of 2-fluoro-4-methyl-1-nitrobenzene (2.22 g, 14.3 mmol) and $Na_2Cr_2O_7$ (5.78g, 20.0 mmol) in water (100 mL), concentrated solution of H_2SO_4 (80 mL) is added dropwise over the period of 1 h. A dark brownish-black solution was stirred overnight at rt. The mixture was then gently refluxed for 2h, cooled to rt, combined with ice cold water (500 mL) then filter. Precipitate was washed several time with water and dried under vacuum to yield 2.05 g (77%) white solid.

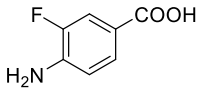


3-fluoro-4-nitorbenzoic acid **13**. ^{1}H NMR ($CDCl_3$): δ 8.27-8.23 (m, 1H, Ar-H(3)), 8.07-8.06 (m, 1H, Ar-H(2)), 8.04 (d, 1H, J = 7.5 Hz, Ar-H(6)), ^{19}F NMR: δ -119.31

4-amino-3-fluorobenzoic acid (14)

Reduction of nitro group of **13** into an amino group of **14** was performed as reported earlier[40]. To a solution of *3-fluoro-4-nitrobenzoic acid* (**13**) (2.0 g, 10.8 mmol), 0.1 mole % of 5% Pd/C and Et_3N (6.34 mL; 45.5 mmol) 97% HCOOH solution (2.5 mL) is added in five portions over the period of 30 min. The reaction mixture was refluxed for 1 h then stirred overnight at rt. Reaction was monitored by TLC, and upon completion the mixture is taken in EtOAc (30 mL), filtered through celite, and washed with H_2O (10 ml x 3),

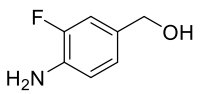
aqueous layers were extracted with EtOAc (20 mL x 5). Organic layers were washed with brine (15 mL x 1) and dried over anhydrous MgSO4 the concentrated over under reduced pressure and product was dried under vacuum to yield 1.23 g (74%) off white solid.



4-amino-3-fluorobenzoic acid **14**. ^{1}H NMR ($CDCl_3$): δ 7.65-7.62 (m, 2H, Ar-H(2,6)), 6.72 (t, 1H, *J* = 7.4 Hz Ar-H(5)), ^{19}F NMR: δ -136.29

(4-amino-3-fluorophenyl)methanol (15)

To an ice cold solution of 4-amino-3-fluorobenzoic acid (**14**; 700.8 mg/4.52 mmol) in dry THF (50 mL), $LiAlH_4$ (188.7 mg; 4.97 mmol) was added in four portions over a period of 20 min. The reaction mixture was stirred 24 overnight at rt under nitrogen gas. The completion of reaction was monitored by TLC. The reaction suspension was filtered through celite pad and washed with THF (2 x 10 mL). Solvent was evaporated under reduced pressure and residue was dissolved in EtOAc (50 mL) and washed with water (1 x 10mL), 15% aqueous NaOH (2 x 10mL), water (1 x 10 mL) and brine (1 x 10mL). The organic layer was dried over anhydrous $MgSO_4$ and concentrated under reduced pressure. The product was isolated from flash silica gel column chromatography using 10% EtOAc/Hexane, and yielded 0.395 g (62%) transparent oil.



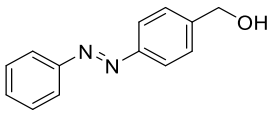
(4-amino-3-fluorophenyldiazynyl)methanol **15**. Yield 45%, ^{1}H NMR (CD_3COCD_3): δ 6.97 (dd *J* = 12.2, 1.6 Hz, 1H, Ar-H), 6.88-6.85 (m, 1H), 6.80 (dd *J* = 9.6, 9.2 Hz, 1H, Ar-H), 4.45 (s, 2H), ^{19}F NMR: δ -137.6

(*E*)-(4-(phenyldiazenyl)phenyl)methanol (9a)

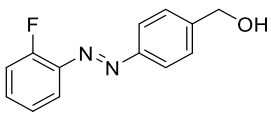
(*E*)-(4-((2-fluorophenyl)diazenyl)phenyl)methanol (9b)

(*E*)-3-fluoro-4-((2-fluorophenyl)diazenyl)-benzoic acid (9c).

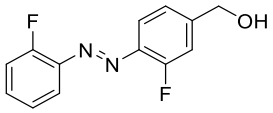
Compound **9a**, **9b** and **9c** were prepared using the procedure described earlier[40]. A solution of 4-aminobenzyl alcohol (**8**) (10 mmol) and nitrosobenzene (**7**) (11 mmol) in glacial acetic acid (50 mL) was stirred overnight at rt to yield product **9a**. Similarly, 4-aminobenzyl alcohol (**8**) and 1-fluoro-2-nitrosobenzene (**11**) were reacted in glacial acetic acid to obtained product **9b**. For product **9c**, 4-amino-3-fluorobenzyl alcohol (**15**) and 1-fluoro-2-nitrosobenzene (**11**) were reacted in glacial acetic acid overnight at rt. The completion of the reaction was monitored by TLC. The solvent was evaporated under reduced pressure. The residue was taken up in EtOAc (50 mL, washed with a saturated aqueous solution of $NaHCO_3$, (2 x 10 mL), and brine (1 x 10 mL). The organic phase was dried over $MgSO_4$. Filtered and concentrated and the product was isolated from flash silica gel column chromatography.



(E)-(4-(phenyldiazenyl)phenyl)methanol **9a**. The product was purified by silica gel flash chromatography (20% EtOAc/Hexane mixtures) to provide a yellow solid 85% yield of **9a**. ^1H NMR (CD_3COCD_3): δ 7.93-7.9 (m, 4H, Ar-H), 7.62-7.52 (m, 5H, Ar-H), 4.76 (d, J = 5.8 Hz, 2H, CH_2)



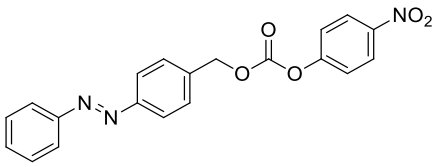
(E)-(4-((2-fluorophenyl)diazenyl)phenyl)methanol **9b**. The product was purified by silica gel flash chromatography (25% EtOAc/Hexane mixtures) to provide a yellow solid 65% yield of **9b**. ^{1}H NMR ($CDCl_3$): δ 7.96 (d, *J* = 6.5 Hz, 2H), 7.76 (dt, *J* = 9.7, 1.7 Hz, 1H, Ar-H), 7.53 (d, *J* = 12.5Hz, 2H), 7.47-7.42 (m, 1H), 7.29-7.20 (m, 2H, Ar-H), 4.79 (s, 2H, CH_2), ^{13}C NMR: δ -161.8, 158.9, 152.4, 144.4, 140.8, 132.6, 132.5, 127.5, 124.4, 123.5, 117.9, 117.3, 64.9, ^{19}F NMR: δ -124.44



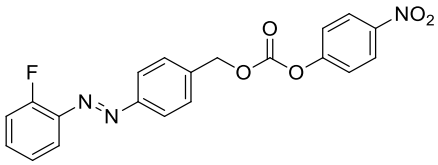
(E)-(3-fluoro-4-((2-fluorophenyl)diazenyl)phenyl)methanol **9c**. The product was purified by silica gel flash chromatography (20% EtOAc/Hexane mixtures) to provide a yellow oil 45% yield of **9c**. ^{1}H NMR (CD_3COCD_3): δ 7.80-7.74 (d and t, *J* = 15.9, 7.7 Hz, 2H, Ar-H), 7.63-7.6 (m, 1H, Ar-H), 7.44-7.39 (m, 2H, Ar-H), 7.78-7.3 (m, 2H, Ar-H), 4.76 (s, 2H, CH_2), ^{19}F NMR: δ -125.66, -125.77

General preparation for 16a-c

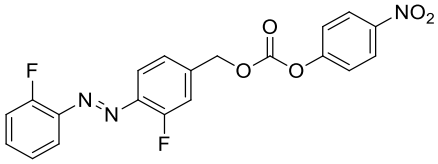
p-nitrochloroformate (0.11 mmol) is added to a solution of azobenzylalcohol analogs (**9a-c**; 0.2 mmol), Et_3N (0.12 mmol) in dry THF (10 mL). The reaction mixture was stirred 24 h at r.t. under nitrogen gas atmosphere. The completion of reaction was monitored by TLC. The solvent was evaporated and residue was dissolved in EtOAc (30 mL) and washed with water (2 x 5 mL), and brine (1 x 5mL). The organic layer was dried over anhydrous $MgSO_4$ and concentrated under reduced pressure. The product was product was isolated from flash silica gel column chromatography.



(E)-4-nitrophenyl 4-(phenyldiazenyl)benzyl carbonate **16a**, The product was purified by silica gel flash chromatography (5% EtOAc/Hexane mixtures) to provide a yellow orange solid 72% yield of **9c**. ^{1}H NMR (CD_3COCD_3): δ 8.36 (d, *J* = 9.5 Hz, 2H), 8.0 (d, *J* = 8.6 Hz, 2H, Ar-H), 7.96 (dd, *J* = 8.3, 1.7 Hz, 2H, Ar-H), 7.73 (dt, *J* = 8.1Hz, 2H, Ar-H), 7.60 (d, *J* = 9.5 Hz, 2H, Ar-H)



(E)-4-((2-fluorophenyl)diazenyl)benzyl (4-nitrophenyl) carbonate **16b,** The product was purified by silica gel flash chromatography (10% EtOAc/Hexane mixtures) to provide a reddish-yellow solid 58% yield of **9c**. ^{1}H NMR ($CDCl_3$): δ 8.28 (d, *J* = 9.1 Hz, 2H), 8.0 (d, *J* = 8.4 Hz, 2H, Ar-H), 7.77 (dt, *J* = 13.8, 1.8 Hz, 1H, Ar-H), 7.6 (d, *J* = 6.6 Hz, 2H, Ar-H), 7.48 (m, 1H), 7.4 (d, *J* = 11.5 Hz, 2H, Ar-H), 7.3-7.23 (m, 2H, Ar-H), 5.37 (s, 2H, CH_2), ESI-MS: m/z: calc. $C_{20}H_{14}FN_3O_5$ [M+H$^+$] 395.1435, found 395.0917

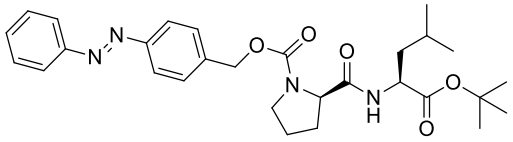


(E)-3-fluoro-4-((2-fluorophenyl)diazenyl)benzyl (4-nitrophenyl) carbonate **16c,** The product was purified by silica gel flash chromatography (15% EtOAc/Hexane mixtures) to provide a yellow solid 30% yield of **16c**. ^{1}H NMR ($CDCl_3$): δ 8.36 (d, *J* = 9.3 Hz, 2H), 7.85-7.78 (m, 2H) 7.68-7.62 (m, 1H, Ar-H), 7.61 (d, *J* = 9.3 Hz, 2H, Ar-H), 7.56-7.53 (m, 1H, Ar-H), 7.49-7.46 (m, 2H, Ar-H), 7.36 (t, *J* = 5.4 Hz, 1H, Ar-H), 5.47 (s, 2H, CH_2), ^{13}C NMR: δ 162.0, 159.9, 156.6, 153.2, 149.9, 146.5, 142.3, 138.2, 134.9, 130.9, 126.1, 125.7, 125.5, 123.9, 122.3, 118.8, 117.3, 69.6, ^{19}F NMR: δ -125.0, 125.4, ESI-MS: m/z: calc. $C_{20}H_{14}FN_3NaO_5$ [M+Na$^+$] 436.0721, found 436.0721

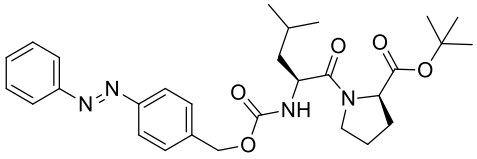
General preparation for (17a-d)

Deprotection of Fmoc- group: Fmoc-amino acid-ester **3**. or **6**. (0.5 mmol) was dissolved into a 20% (v/v) of Diethyl amine (Et_2NH) solution in CH_3CN (10 mL). The reaction mixture was stirred at for 30 min at rt. The solvent was evaporated under reduced pressure and the crude product was used without further purification.

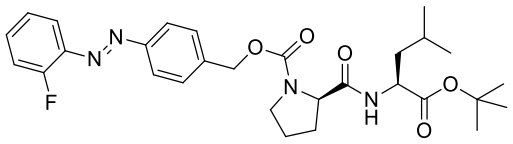
Coupling of amino acid: Unpurified H_2N-Leu-OtBut or H_2N-Pro-OtBut (0.5 mmol) was dissolved in dry DCM (10 mL) and then Et3N (0.55 mmol) was added. Then compound **16a-c** was added into above mixture. The reaction mixture was stirred under nitrogen gas atmosphere for 16 hours. Completion of reaction was monitored by TLC. Solvent was evaporated under reduced pressure and crude was dissolved in EtOAc (25 mL) and organic layer was washed with water (5 mL) and brine (5 mL). The organic layers was dried with anhydrous MgSO4 and concentrated in vacuo. The product **17a-d** was isolated by flash chromatography



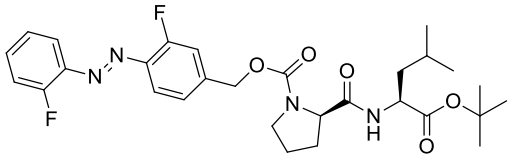
(S)-4-((E)-phenyldiazenyl)benzyl 2-(((S)-1-(tert-butoxy)-4-methyl-1-oxopentan-2-yl)carbamoyl)pyrrolidine-1-carboxylate **17a,** The product was purified by silica gel flash chromatography (25% EtOAc/Hexane mixtures) to provide a yellow solid 37% yield of **17a**.
ESI-MS: m/z: calc. $C_{29}H_{38}N_4O_4$ [M^+] 522.2842, found 522.9815



(S)-tert-butyl 1-((S)-4-methyl-2-((((4-((E)-phenyldiazenyl)benzyl)oxy)carbonyl)amino)penta-noyl)pyrrolidine-2-carboxylate **17b,** The product was purified by silica gel flash chromatography (25% EtOAc/Hexane mixtures) to provide a yellow solid 40% yield of **17b** ESI-MS: m/z: calc. $C_{29}H_{38}N_4NaO_5$ [$M+Na^+$] 545.2740, found 545.2992



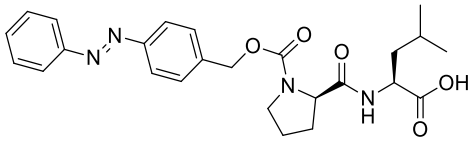
(S)-4-((E)-(2-fluorophenyl)diazenyl)benzyl 2-(((S)-1-(tert-butoxy)-4-methyl-1-oxopentan-2-yl)carbamoyl)pyrrolidine-1-carboxylate **17c,** The product was purified by silica gel flash chromatography (25% EtOAc/Hexane mixtures) to provide a yellow solid 66% yield of **17c,** ESI-MS: m/z: calc. $C_{29}H_{37}FN_2NaO_5$ [M+Na^+] 563.2646, found 563.3235



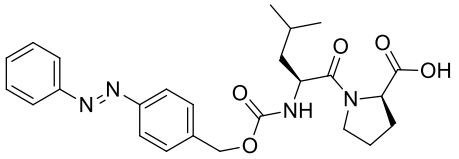
(S)-3-fluoro-4-((E)-(2-fluorophenyl)diazenyl)benzyl 2-(((S)-1-(tert-butoxy)-4-methyl-1-oxopentan-2-yl)carbamoyl)pyrrolidine-1-carboxylate **17d,** The product was purified by silica gel flash chromatography (30% EtOAc/Hexane mixtures) to provide a yellow solid 61% yield of **17d,** ESI-MS: m/z: calc. $C_{29}H_{36}N_2F_4NaO_5$ [M+Na^+] 581.2551, found 581.2929

General preparation for (18a-d)

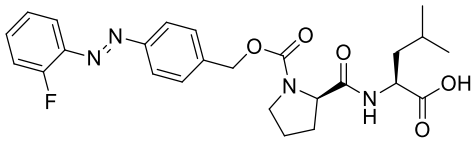
Carboxylic protecting (*tert*-Butyl) group from compound **17a-d** was deprotected by refluxing compound **17a-d** (0.1 mmol) and *p*-toluenesulfonic acid (0.06 mmol) in benzene (3 mL) under anhydrous conditions and completion of reaction was monitored by TLC. Deprotection complete in 3 hours. Then the solvent was evaporated under reduced pressure and product was isolated from crude by flash chromatography



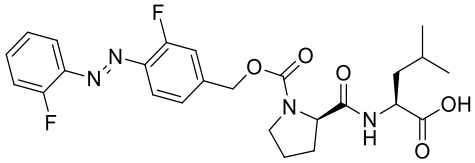
(S)-4-methyl-2-((S)-1-(((4-((E)-phenyldiazenyl)benzyl)oxy)carbonyl)pyrrolidine-2-carboxamido)pentanoic acid (ABDP) **18a,** The product was purified by silica gel flash chromatography (1% MeOH/$CHCl_3$ mixtures) to provide a yellow solid 87% yield of **18a**, ESI-MS: m/z: calc. $C_{25}H_{31}N_4$ NaO_5 [M+Na^+] 489.2114, found 489.2285



(S)-tert-butyl 1-((S)-2-((((3-fluoro-4-((E)-(2-fluorophenyl)diazenyl)benzyl)oxy)carbonyl)amino)-4-methylpentanoyl)pyrrolidine-2-carboxylate (RABDP) **18b,** The product was purified by silica gel flash chromatography (1% MeOH/$CHCl_3$ mixtures) to provide a yellow solid 75% yield of **18b**, ESI-MS: m/z: calc. $C_{25}H_{31}N_4\ NaO_5$ $[M+Na^+]$ 489.2114, found 489.2285



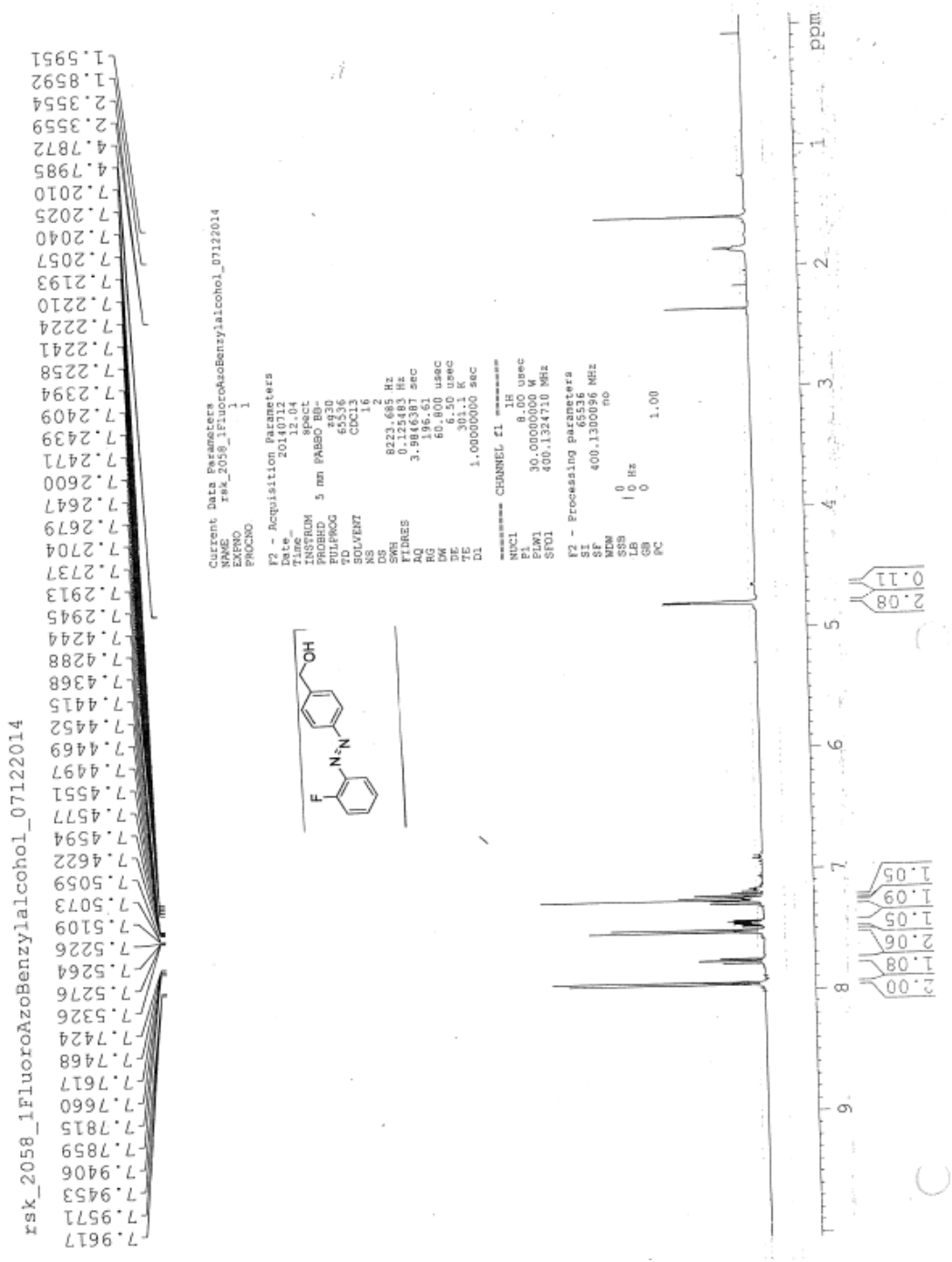
(S)-2-((S)-1-(((4-((E)-(2-fluorophenyl)diazenyl)benzyl)oxy)carbonyl)pyrrolidine-2-carboxamido)-4-methylpentanoic acid (1FABDP) **18c,** The product was purified by silica gel flash chromatography (2% MeOH/$CHCl_3$ mixtures) to provide a yellow solid 30% yield of **18c**, ^{19}F NMR: δ -124.3, ESI-MS: m/z: calc. $C_{25}H_{29}FN_4NaO_5$ $[M+Na^+]$ 507.2020, found 507.2658



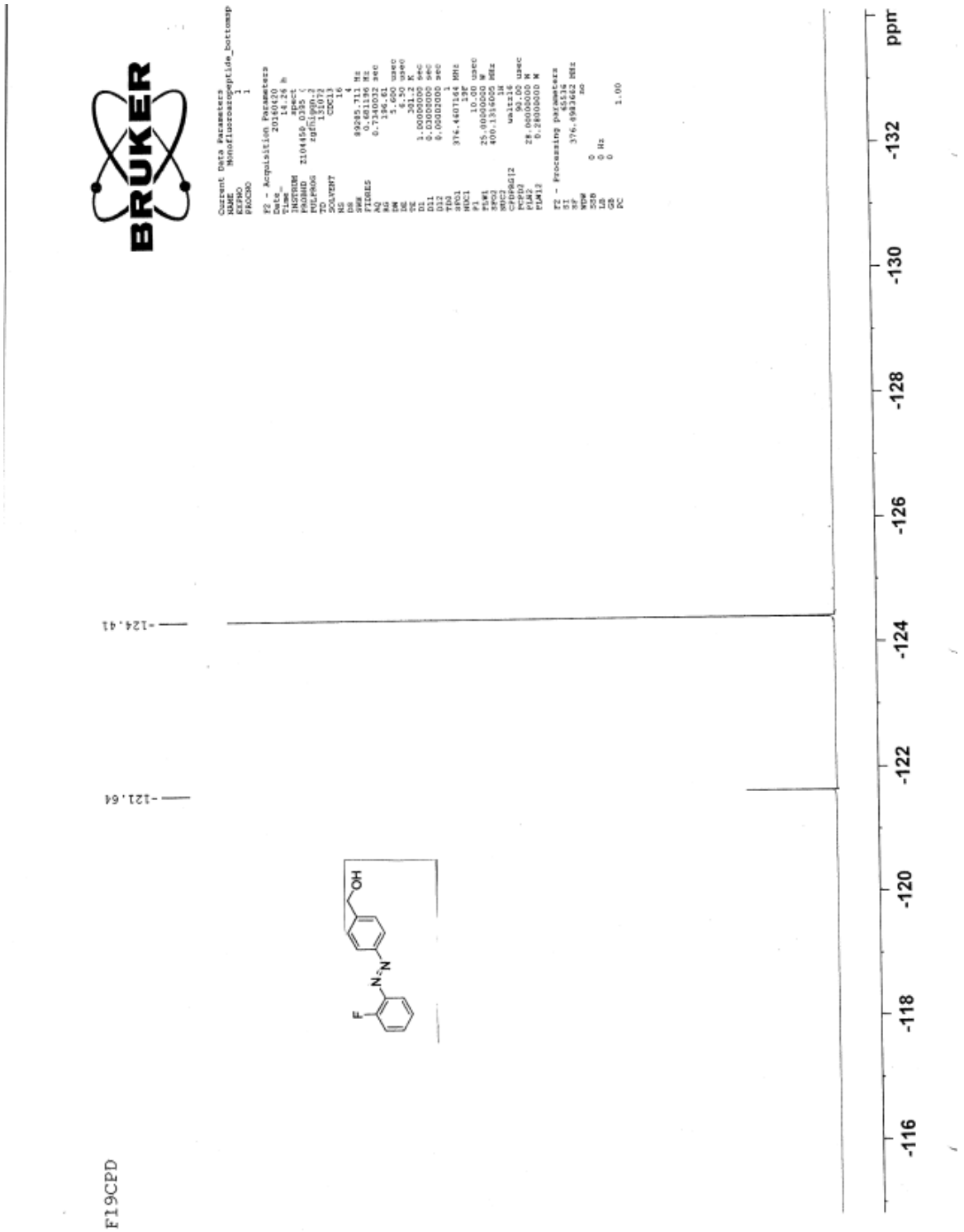
(S)-2-((S)-1-(((3-fluoro-4-((E)-(2-fluorophenyl)diazenyl)benzyl)oxy)carbonyl)pyrrolidine-2-carboxamido)-4-methylpentanoic acid (2FABDP) **18d,** The product was purified by silica gel flash chromatography (2% MeOH/$CHCl_3$ mixtures) to provide a yellow solid 28% yield of **18d,** ESI-MS: m/z: calc. $C_{25}H_{28}F_2N_4NaO_5$ $[M+Na^+]$ 525.1925, found 525.2273

B. NMR Spectra of Synthetic Intermediates and Products

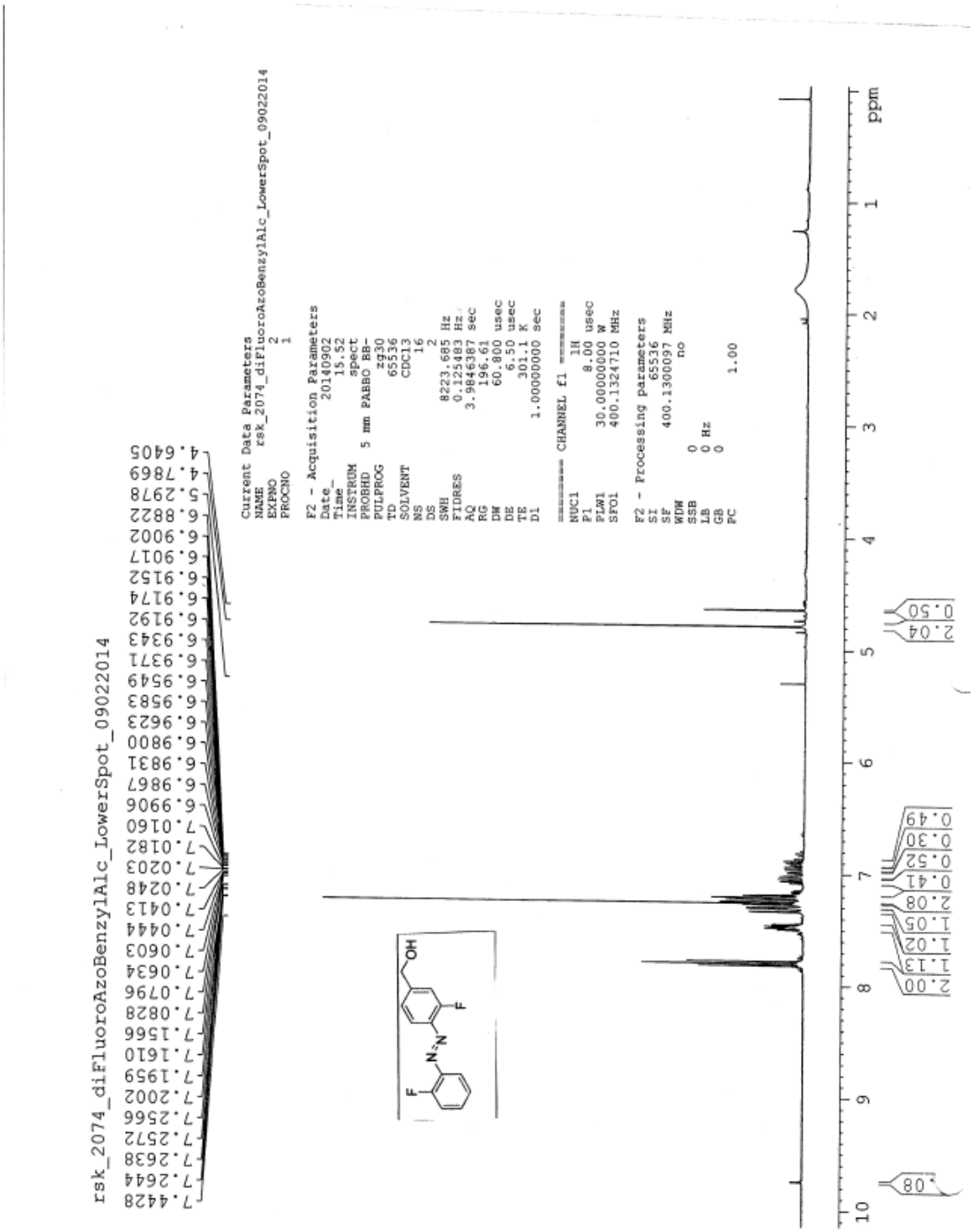
^{1}H NMR of *(E)-(4-((2-fluorophenyl)diazenyl)phenyl)methanol* **9b**.

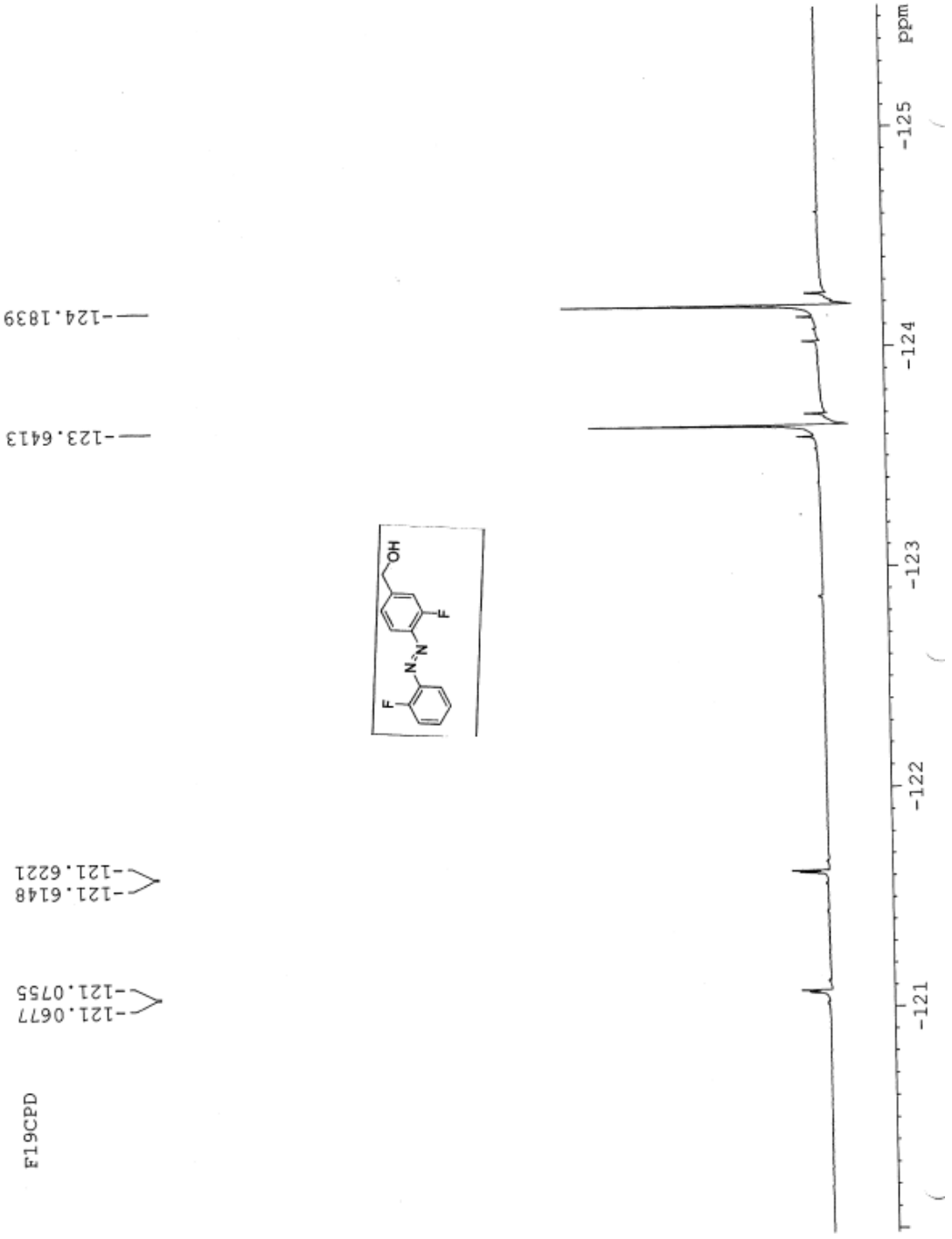


^{19}F NMR of *(E)-(4-((2-fluorophenyl)diazenyl)phenyl)methanol* **9b**.



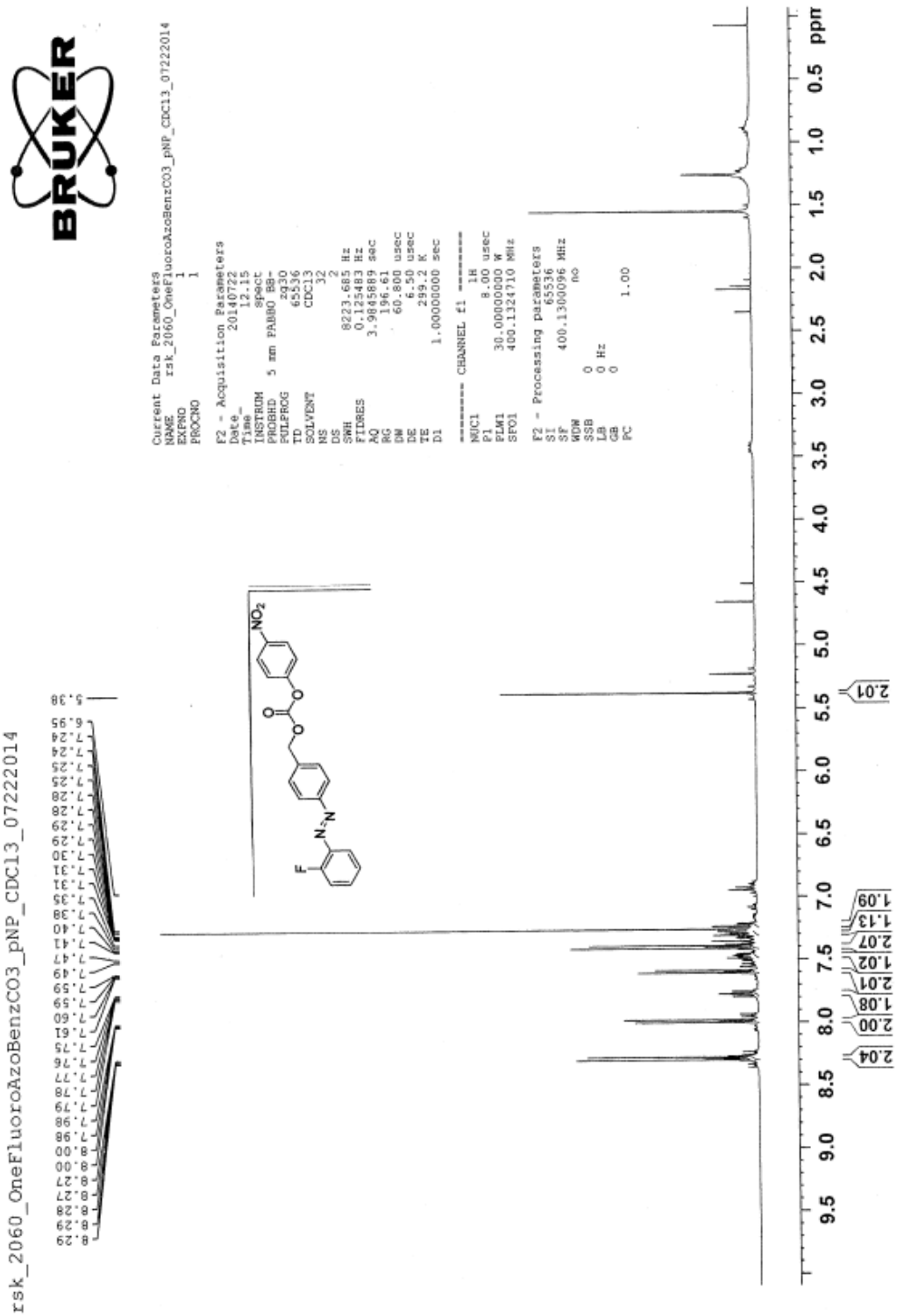
^{1}H NMR of *(E)-(3-fluoro-4-((2-fluorophenyl)diazenyl)phenyl)methanol* **9c**



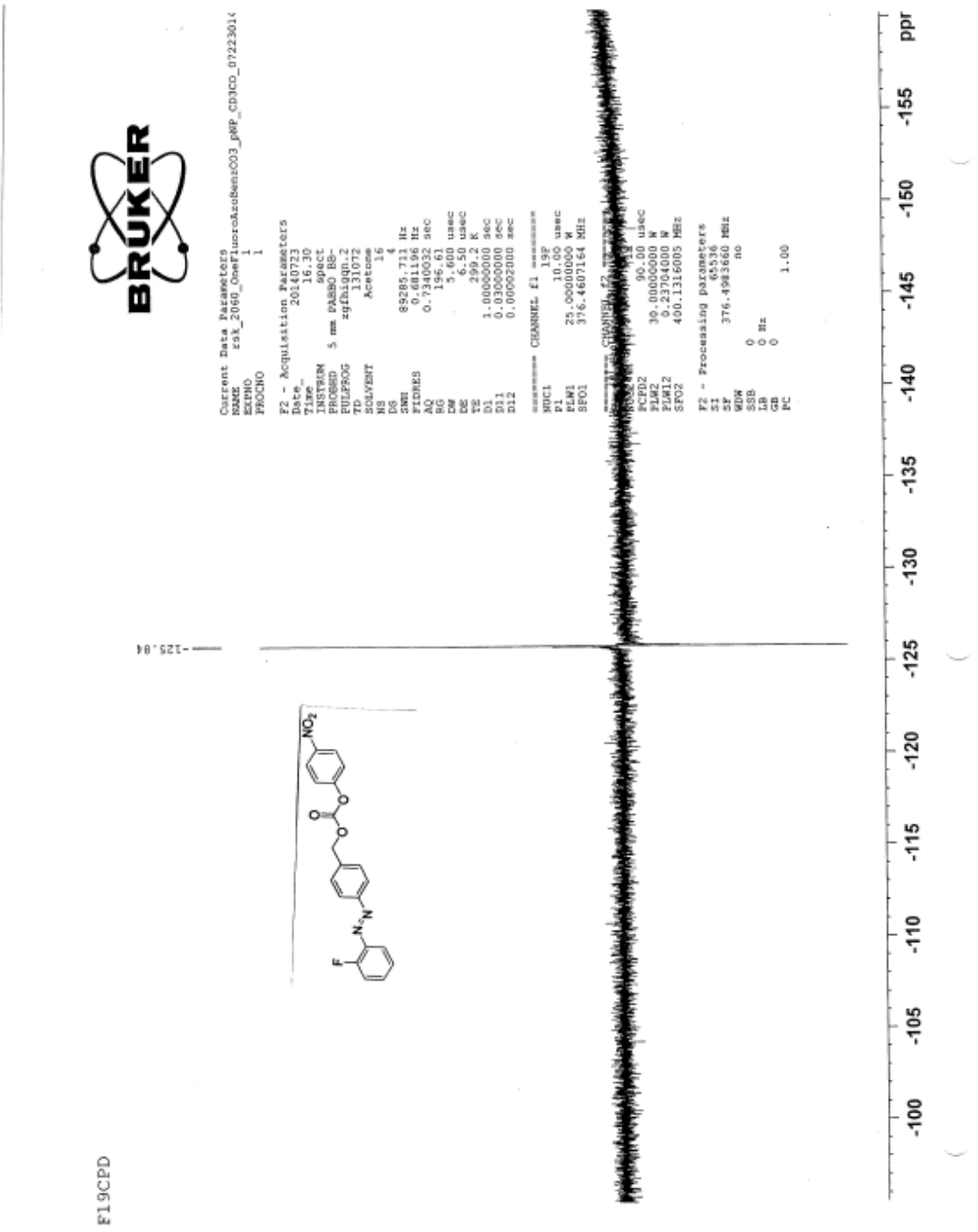


^{19}F NMR of *(E)-(3-fluoro-4-((2-fluorophenyl)diazenyl)phenyl)methanol* **9c**

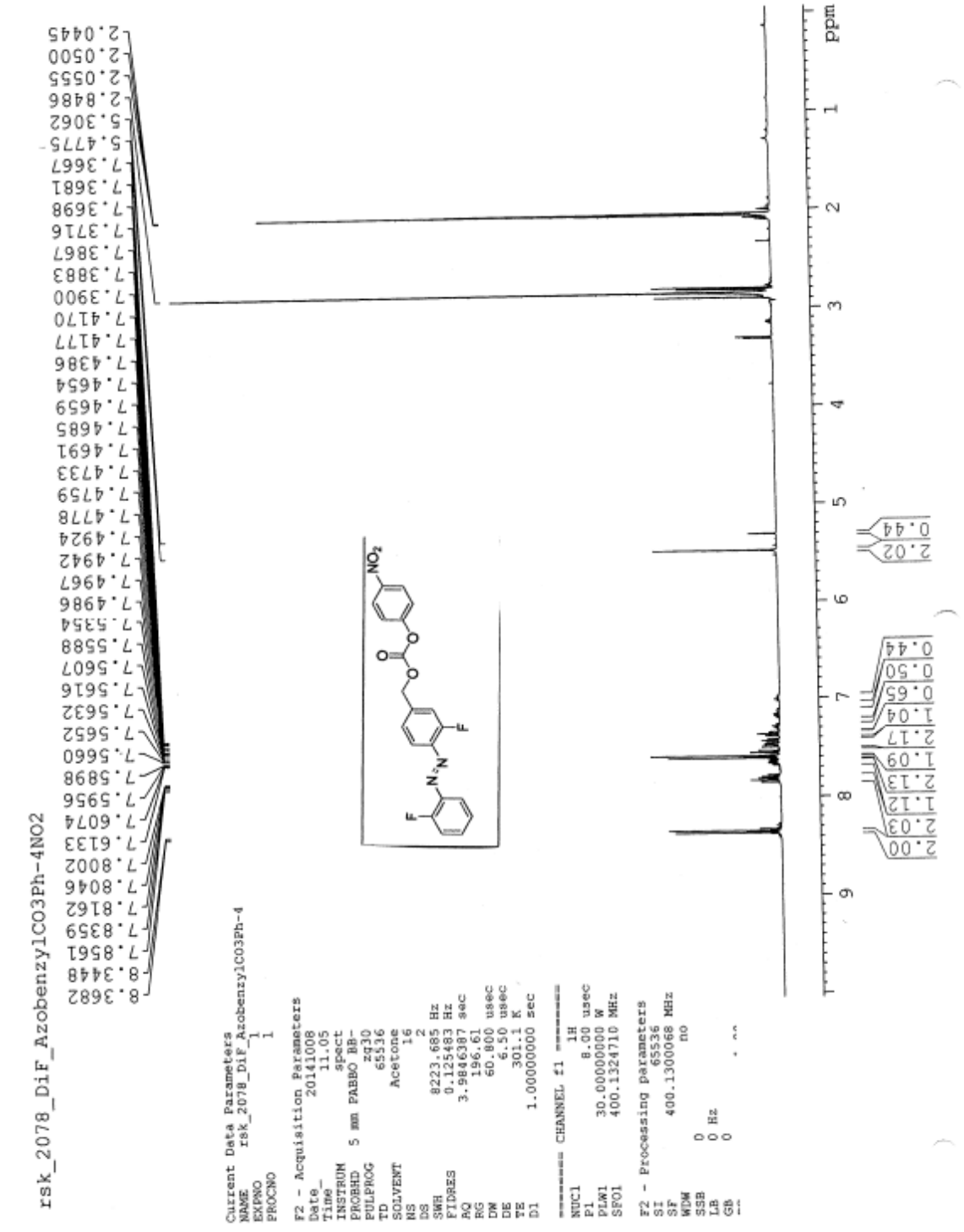
[1]H NMR of *(E)-4-((2-fluorophenyl)diazenyl)benzyl (4-nitrophenyl) carbonate* **16b**



^{19}F NMR of *(E)-4-((2-fluorophenyl)diazenyl)benzyl (4-nitrophenyl) carbonate* **16b**

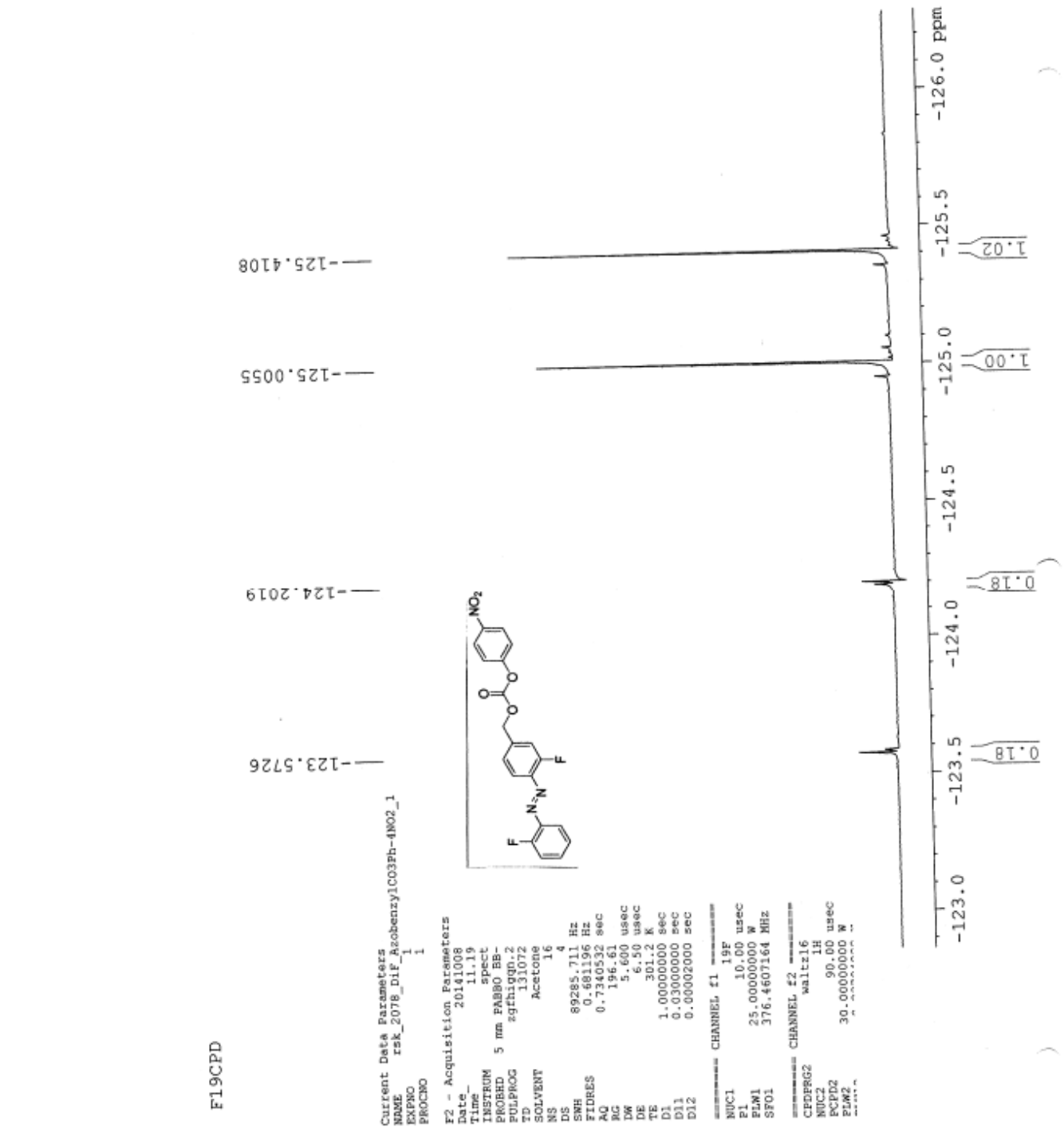


^{1}H NMR of *(E)-3-fluoro-4-((2-fluorophenyl)diazenyl)benzyl (4-nitrophenyl) carbonate*



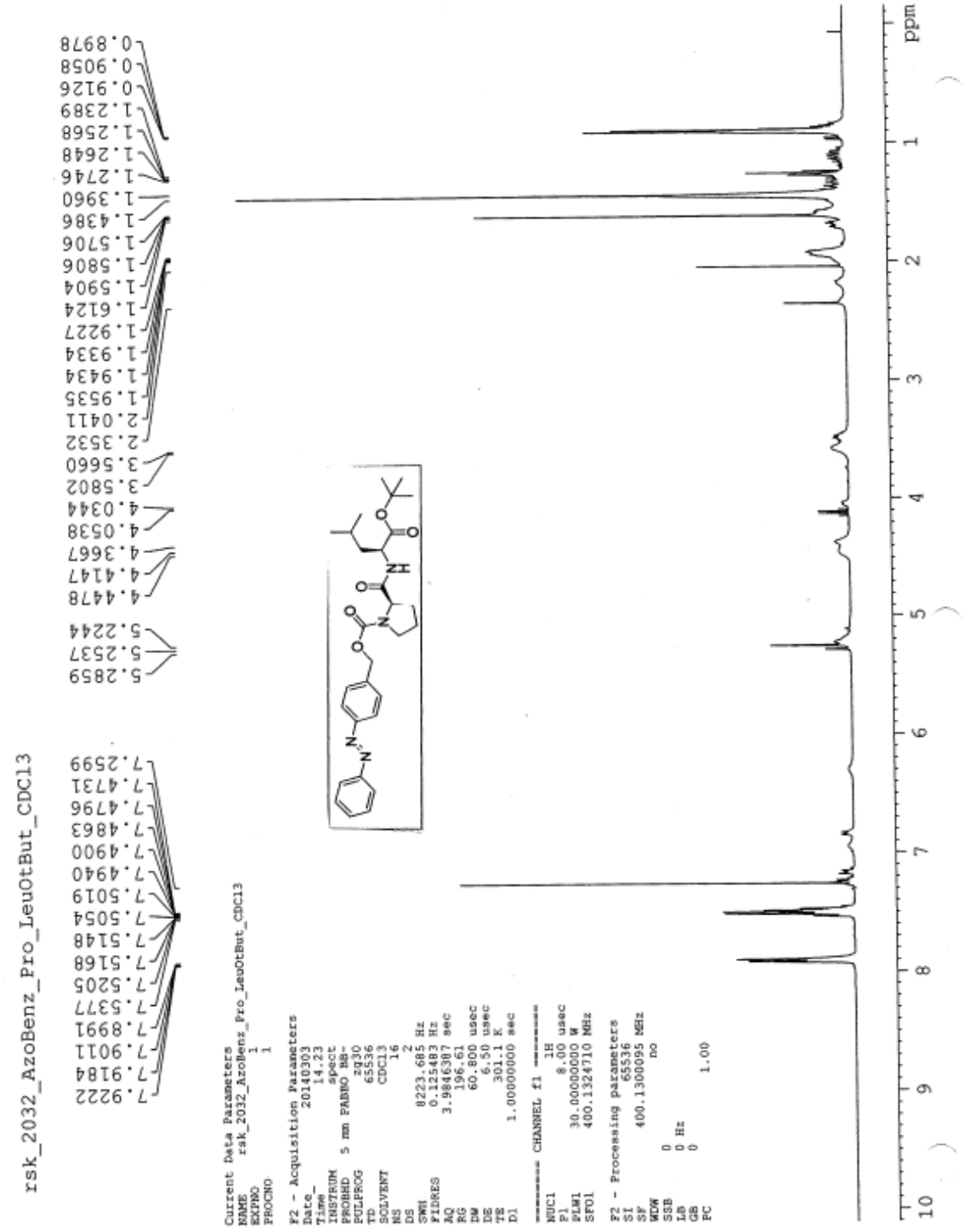
16c

^{19}F NMR of *(E)-3-fluoro-4-((2-fluorophenyl)diazenyl)benzyl (4-nitrophenyl) carbonate*



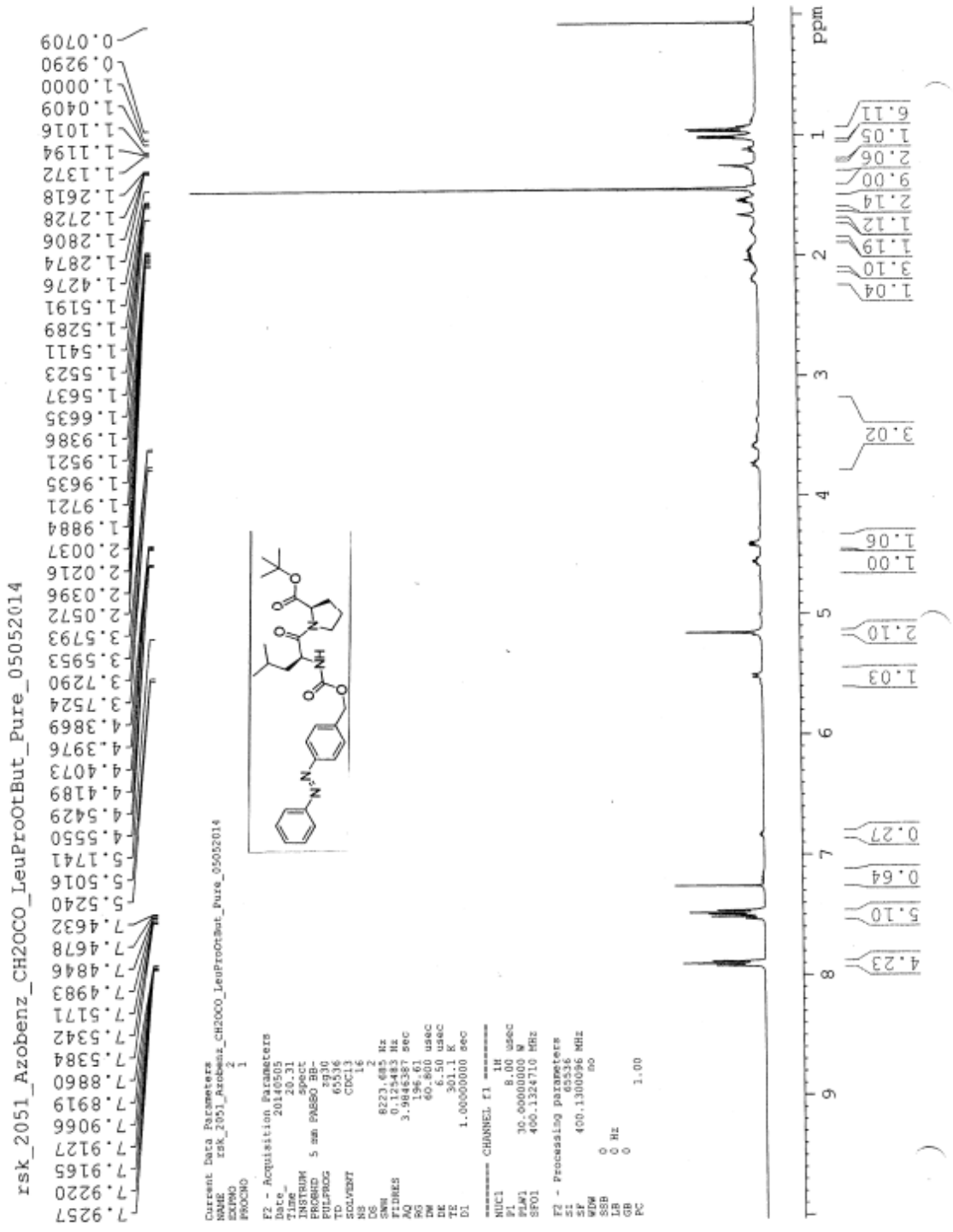
16c

^{1}H NMR of *(S)-4-((E)-phenyldiazenyl)benzyl 2-(((S)-1-(tert-butoxy)-4-methyl-1-*



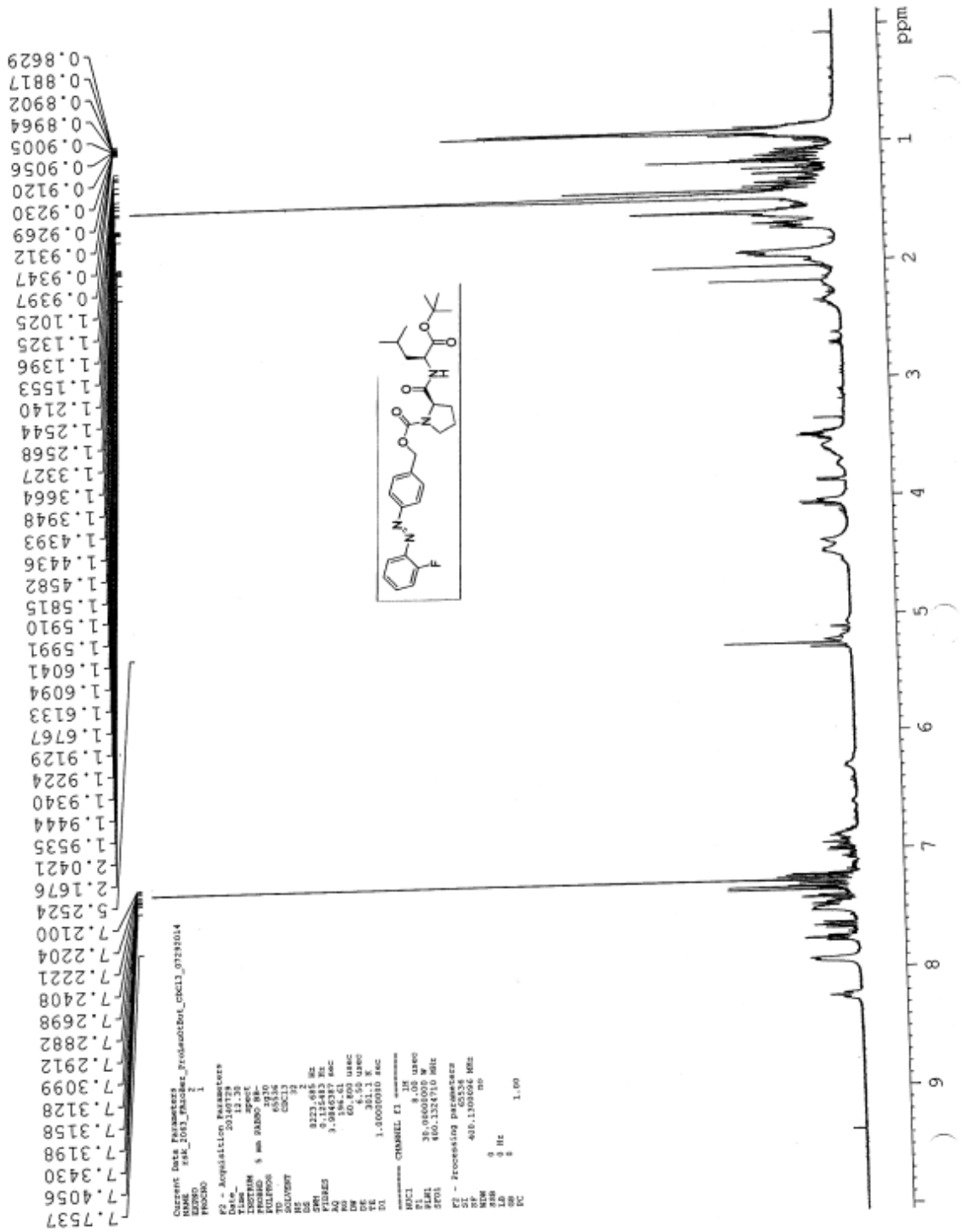
oxopentan-2-yl)carbamoyl)pyrrolidine-1-carboxylate **17a**

^{1}H NMR of *(S)-tert-butyl 1-((S)-4-methyl-2-((((4-((E)-*



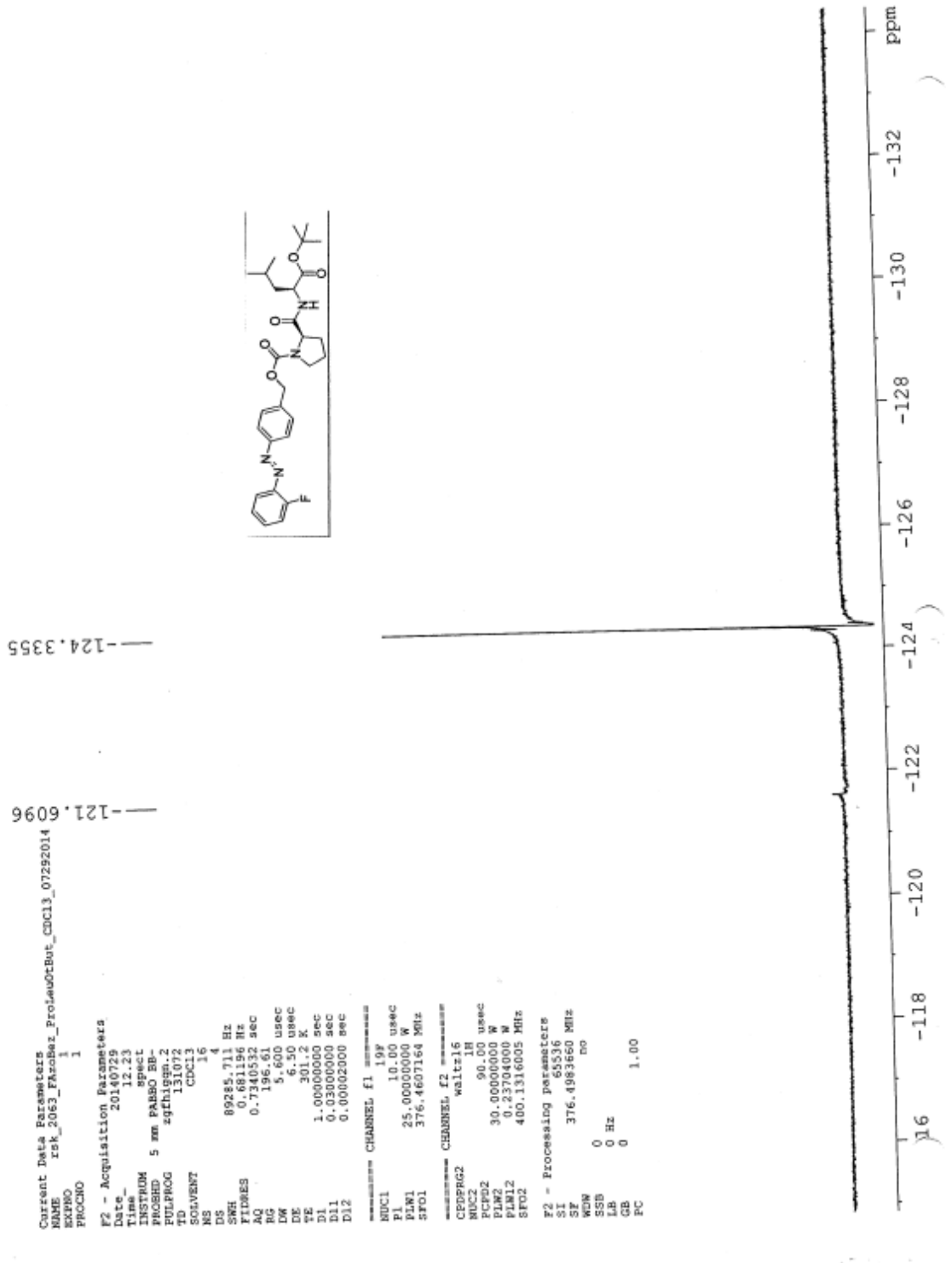
phenyldiazenyl)benzyl)oxy)carbonyl)amino)penta-noyl)pyrrolidine-2-carboxylate **17b**

^{1}H NMR of *(S)-4-((E)-(2-fluorophenyl)diazenyl)benzyl 2-(((S)-1-(tert-butoxy)-4-methyl-*



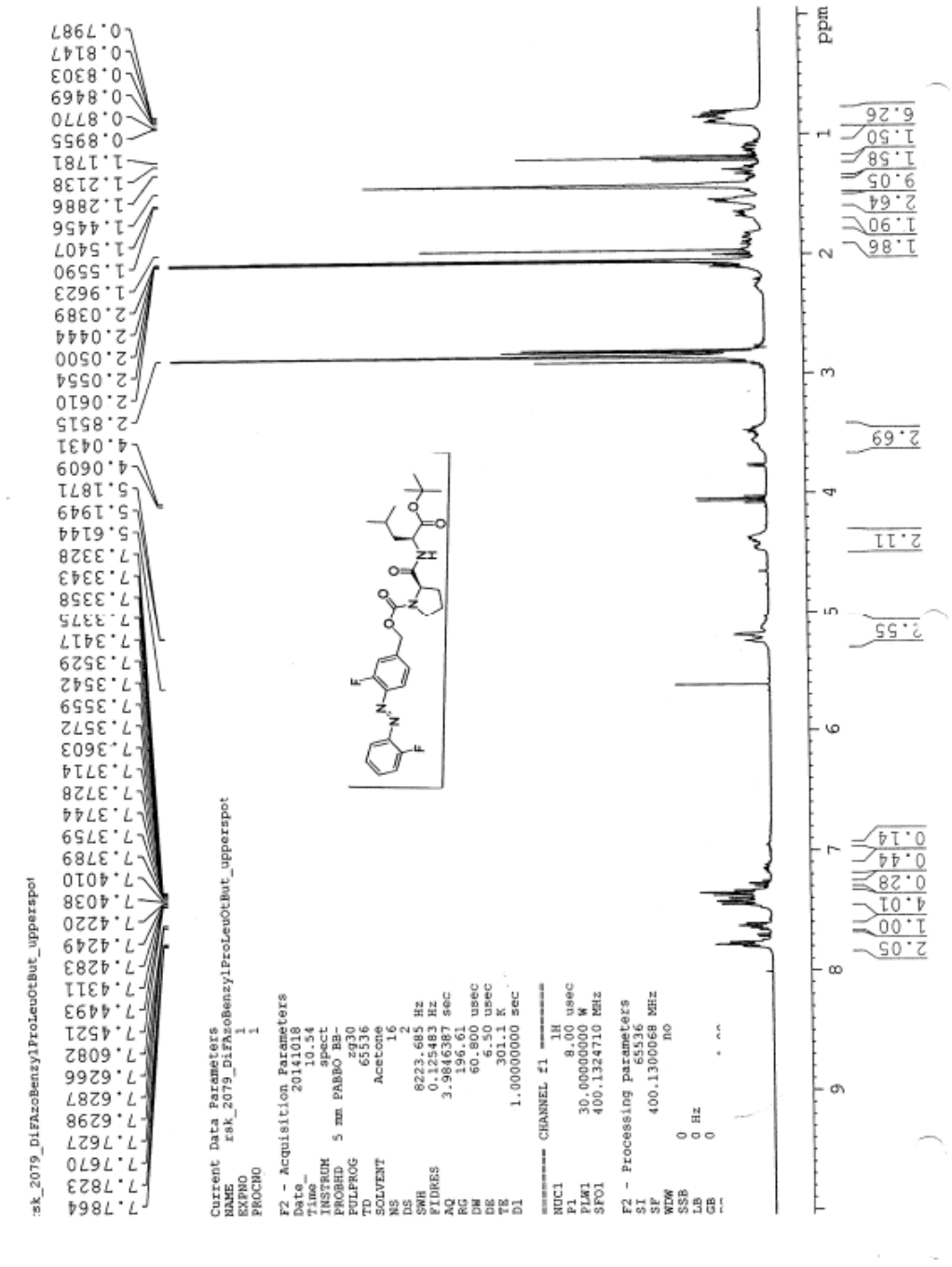
1-oxopentan-2-yl)carbamoyl)pyrrolidine-1-carboxylate **17c**

^{1}H NMR of *(S)-4-((E)-(2-fluorophenyl)diazenyl)benzyl 2-(((S)-1-(tert-butoxy)-4-methyl-*

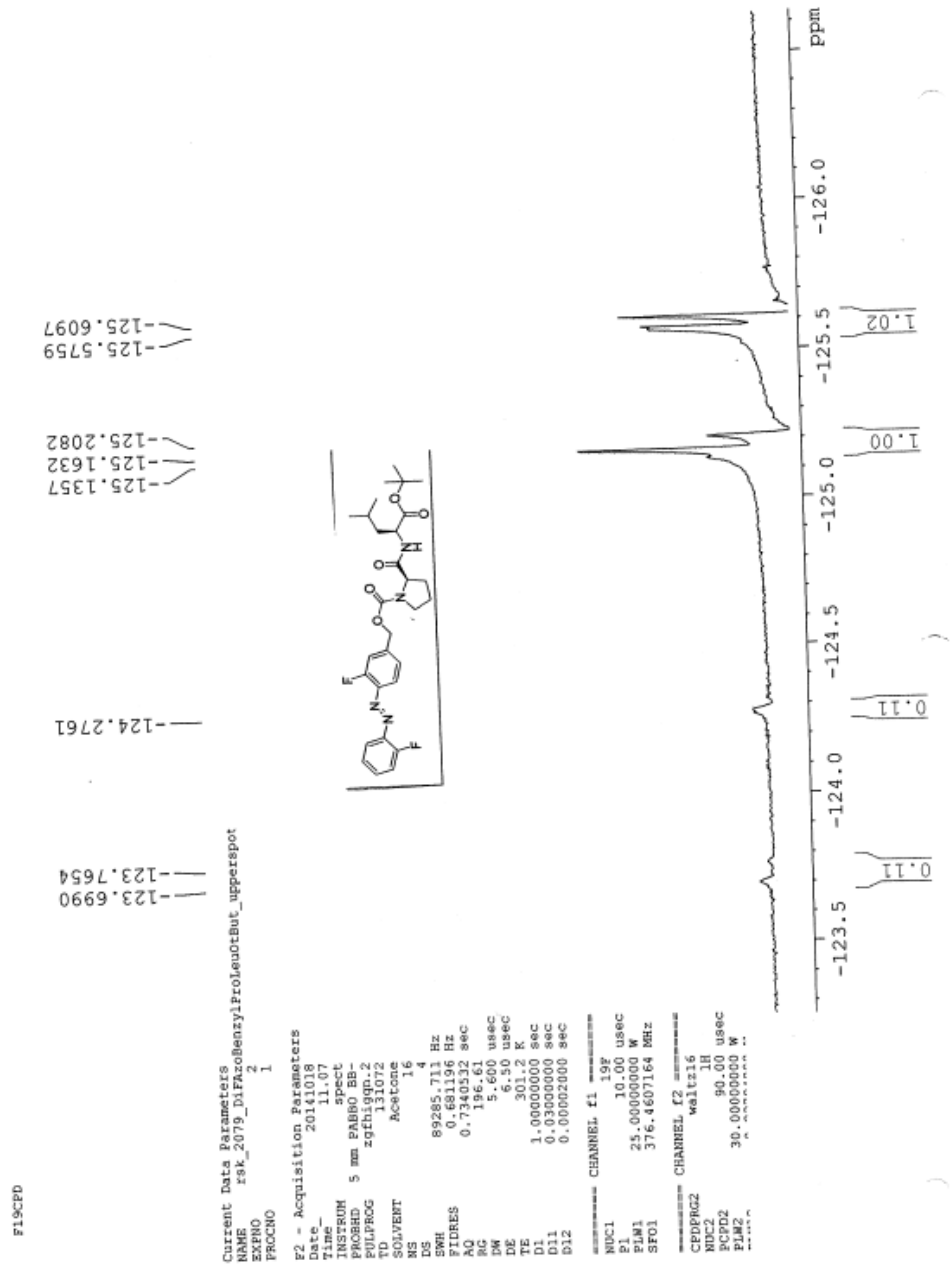


1-oxopentan-2-yl)carbamoyl)pyrrolidine-1-carboxylate **17c**

^{1}H NMR of *(S)-2-((S)-1-(((3-fluoro-4-((E)-(2-fluorophenyl)diazenyl)benzyl)oxy)carbonyl)pyrrolidine-2-carboxamido)-4-methylpentanoic acid* (2FABDP) **18d**



^{19}F NMR of *(S)-2-((S)-1-(((3-fluoro-4-((E)-(2-fluorophenyl)diazenyl)benzyl)oxy)carbonyl)pyrrolidine-2-carboxamido)-4-methylpentanoic acid*



REFERENCES

1. Russew, M. M.; Hecht, S. Photoswitches: From Molecules to Materials. *Adv. Mater.* **2010**, *22* (31), 3348–3360 DOI: 10.1002/adma.200904102.

2. Sheng, Y.; Chen, Q.; Yao, J.; Wang, Y.; Liu, H. Hierarchical Assembly of a Dual-responsive Macroscopic Insulated Molecular Wire Bundle in a Gradient System. *Sci. Rep.* **2015**, *5* (Cd), 7791 DOI: 10.1038/srep07791.

3. Merino, E.; Ribagorda, M. Control Over Molecular Motion Using the Cis-trans Photoisomerization of the Azo Group. *Beilstein J. Org. Chem.* **2012**, *8*, 1071–1090 DOI: 10.3762/bjoc.8.119.

4. Kumar, A. S.; Ye, T.; Takami, T.; Yu, B. C.; Flatt, A. K.; Tour, J. M.; Weiss, P. S. Reversible Photo-switching of Single Azobenzene Molecules in Controlled Nanoscale Environments. *Nano Lett.* **2008**, *8* (6), 1644–1648 DOI: 10.1021/nl080323+.

5. Hartley, G. S. The Cis-form of Azobenzene. *Nature*. **1937**, 281–281.

6. A Stable cis-Azobenzene in Aqueous Solution. Chambers Haworth

7. Wagner, N.; Theato, P. Light-induced Wettability Changes on Polymer Surfaces. *Polym. (United Kingdom)* **2014**, *55* (16), 3436–3453 DOI: 10.1016/j.polymer.2014.05.033.

8.Willem A. V,; Szymanski, V.; and Feringa, B. L.; Photopharmacology: Beyond Proof of Principle. *J. Am. Chem. Soc.* **2014**, 136 (6), pp 2178–2191 DOI: 10.1021/ja413063e

9. Ayouni, L.; Cazorla, G.; Chaillou, D.; Herbreteau, B.; Rudaz, S.; Lantéri, P.; Carrupt, P. -a. Fast Determination of Lipophilicity by HPLC. *Chromatographia* **2005**, *62* (5-6), 251–255 DOI: 10.1365/s10337-005-0608-6.

10. Terada, H. Determination of Log P_{oct} by High-Performance Liquid Chromatography, and its Application in the Study of Quantitative Structure-Activity Relationships. *Quant. Struct. Relationships* **1986**, *5* (3), 81–88 DOI: 10.1002/qsar.19860050302.

11. Morales, I.; Frederick, J. PZ-Peptidase from Chick Embryos. **1977**, *252* (14), 4855–4860.

12. Knie, C.; Utecht, M.; Zhao, F.; Kulla, H.; Kovalenko, S.; Brouwer, a M.; Saalfrank, P.; Hecht, S.; Bleger, D. ortho-Fluoroazobenzenes: Visible Light Switches with Very Long-Lived Z Isomers. *Chemistry (Easton).* **2014**, *20* (50), 16492–16501 DOI: 10.1002/chem.201404649.

13. Yoon, J. H.; Yoon, S. Photoisomerization of Azobenzene Derivatives Confined in Gold Nanoparticle Aggregates. *Phys. Chem. Chem. Phys.* **2011**, *13* (28), 12900–12905 DOI: 10.1039/c0cp02588g.

14. Brode, W.; Gould, J.; Wyman, G. The Relation between the Absorption Spectra and the Chemical Constitution of Dyes. XXV. Phototropism and cis-trans Isomerism in Aromatic Azo Compounds1. *J. Am. Chem. Soc.* **1952**, *74* (9), 4641–4646 DOI: 10.1021/ja01138a059.

15. Sangster, J.; Octanol-Water Partition Coefficients of Simple Organic Compounds. *Journal of Physical and Chemical Reference Data*. **1989**, 1111–1229.

16. Hansch, C.; A Quantitative Approach to Biochemical Structure-Activity Relationships. *Chem. Res.* **1969**, *2* (4), 232–239 DOI: 10.1021/ar50020a002.

17. Robinson, J.R.; Lee, V.H. *Controlled Drug Delivery Fundamentals and Applications*, CRC Press, 1987; pp 3-94

18. Du, C. M.; Valko, K.; Bevan, C.; Reynolds, D.; Abraham, M. H. Rapid Method for Estimating Octanol-water Partition Coefficient (Log P Oct) From Isocratic RP-HPLC and a Hydrogen Bond Acidity Term (A). *J. Liq. Chromatogr.* **2001**, 24 (5), 635–649 DOI: 10.1081/JLC-100103400.

19. Abraham, R. J.; Griffiths, L.; Perez, M. 1H NMR spectra. Part 30: 1H Chemical Shifts in Amides and the Magnetic Anisotropy, Electric field and Steric Effects of the Amide Group. *Magn. Reson. Chem*. **2013**, 51 (3), 143–155 DOI: 10.1002/mrc.3920.

20. Shoulders, H.; Welch, S. A Very Brief, Rapid, Simple, and Unified Method for estimating Carbon-13 NMR chemical shifts: the BS method. J. Chem. Educ. **1987**, 64 (11), 915–918.

21. Wedemeyer, W. J.; Welker, E.; Scheraga, H. a. Current Topics Proline Cis - Trans Isomerization and Protein Folding †. Biochemistry **2002**, 41, 14637–14644 DOI: 10.1021/bi020574b.

22. Abraham, R. J.; McLauchlan, K. a. The Proton Resonance Spectra and Conformations of the Prolines. *Mol. Phys.* **1962**, *5* (2), 195–203 DOI: 10.1080/00268976200100201.

23. P. Mukherjee and K. J. Mysels, Critical Micelle Concentration of Aqueous Surfactant Systems. NSRDS-NBS 36, US Government Printing Office, Washington, DC, USA, 1971.

24. Turro, N.J.; Yekta, A. Luminescent Probes for Detergent Solutions. A Simple Procedure for Determination of the Mean Aggregation Number of Micelles. **1978**, *J. Am. Chem. Soc.* 100: 5951. DOI: 10.1021/ja00486a062.

25. Ungnade, H. E.; The Effect of Solvent on the Absorption Spectra of Aromatic Compounds. *J. Am. Chem. Soc.* **1953**, 75 (2) 432-434 DOI: 10.1021/ja01098a051.

26. Bandara, H. M. D.; Burdette, S. C. Photoisomerization in Different Classes of Azobenzene. *Chem. Soc. Rev.* **2012**, 41 (5), 1809–1825 DOI: 10.1039/C1CS15179G.

27. Knie, C.; Utecht, M.; Zhao, F.; Kulla, H.; Kovalenko, S.; Brouwer, a M.; Saalfrank, P.; Hecht, S.; Bleger, D. ortho-Fluoroazobenzenes: Visible Light Switches with Very Long-Lived Z Isomers. *Chemistry (Easton).* **2014**, *20* (50), 16492–16501 DOI: 10.1002/chem.201404649.

28. Merino, E.; Ribagorda, M. Control Over Molecular Motion Using the cis-trans Photoisomerization of the Azo Group. *Beilstein J. Org. Chem.* **2012**, *8*, 1071–1090 DOI: 10.3762/bjoc.8.119.

29. Platts, J.; Abraham, M.; Butina, D.; Hersey, a. Estimation of Molecular Linear Free Energy Relationship Descriptors by a Group Contribution Approach. 2. Prediction of Partition Coefficients. *J. Chem. Inf. Comput. Sci.* **2000**, *40* (1), 71–80 DOI: doi: 10.1021/ci990427t.

30. Abraham, M. H.; Abraham, R. J.; Acree, W. E.; Aliev, A. E.; Leo, A. J.; Whaley, W. L. An NMR Method for the Quantitative Assessment of Intramolecular Hydrogen Bonding; Application to Physicochemical, Environmental, and Biochemical Properties. *J Org Chem*, **2014**, *11083* (1).

31. P.R. Wells. Linear Free-Energy Relationships. AP London-New York 1968

32. Jacobson, J.; Frenz, J.; Horvath, C. Measurement of Adsorption Isotherms by Liquid Chromatography. *J. Chrom. A.* **1984**, 316: 53-68.

33. Tanford, C. of Amino Guanidine - I. *J. Biol. Chem.* **1970**, No. 7, 1648–1652.

34. Faller, B.; Ertl, P. Computational Approaches to Determine Drug Solubility. *Adv. Drug Deliv. Rev.* **2007**, *59* (7), 533–545 DOI: 10.1016/j.addr.2007.05.005.

35. Abbaspourrad, A.; Carroll, N. J.; Kim, S. H.; Weitz, D. Polymer Microcapsules with Programmable Active Release. *J. Am. Chem. Soc.* **2013**, *135* (20), 7744–7750 DOI: 10.1021/ja401960f.

36. Zhi, D.; Guozhong, Y.; Pittman, C. U.; Peng-Yu, Y.; Kai, L.; Mun Hong, N.; Lear, M. J; Wenk, M. R.; Yao, S. Q. Solution-phase Synthesis and Evaluation of Tetraproline Chiral Stationary Phases, *Chirality,* **2012**, *24*(4), 329-38

37. Daniels, N. R.; Melancon, B. J.; Wang, E. A.; Crews, B. C.; Marnett, L. J.; Solukowski, G. A.; Lindsley, C. W. Progress Towards the Total Synthesis of Lucetamycin A: Total Synthesis and Biological Evaluation of 8-*epi*-Lucetamycin A, *J. Org. Chem.,* **2009**, 74, 8852-8855

38. Stöber, P.; Schelhaas, M.; Nägele, E.; Hagenbuch, P.; Rétey, J.; Waldmann, H. Synthesis of Charactristic Lipopeptides of the Human *N*-Ras Protein and their Evaluation as Possible Inhibitors of Protein Farnesyl Transferase, *Bioorg. Med. Chem.*, **1997**, *5*(1), 75-83.

39. Jankowiak, A.; Obijalska, E.; Kaszynski, P.; Pieczonka, A.; Young, V. G. Synthesis and Structural, Spectroscopic, and Electrochemical Characterization of Benzo[c]quinolizinium and its 5-aza-, 6-aza, and 5,6-diaza analogues, *Tetrahedron*, **2011**, *67*, 3317-3327

40. Kamm, O.; Matthews, A. O., *Org. Synth. Coll.*, **1922**, vol *1 392;* **1922**, *vol 2, 53*

41. Stawski, Philipp; Sumser, Martin and Trauner, Dirk, A Photochromic Agonist of AMPA Receptors, *Angew. Chem.* **2012**, *51*(23), 5748-5751

42. Mangravite, John A., Palladium Catalyzed Reduction of Nitrobenezene, *J. Chem. Ed.*, **1983**, *60 439*