

Thesis for the Master's
degree in chemistry

Alexey Bryukhovskiy

**Design and synthesis
of potential inhibitors
of the transcription
factor c-Myb**

60 study points

DEPARTMENT OF CHEMISTRY
Faculty of mathematics and natural
sciences
UNIVERSITY OF OSLO 10/2011



Acknowledgements

The work for this thesis has been done at the Department of Chemistry, University of Oslo.

I give my sincere gratitude to my supervisor Tore Hansen, who not only served as my supervisor but also encouraged and challenged me throughout my academic program. I want to thank him for teaching me a tremendous amount of chemistry and for always listening to my ideas.

This work could not have been done without advices from Gunnar Herstad. I also wish to thank Åsmund Kaupang, Martin Johansen Hennum, Vitthal Narayan Yadav and Kim Alex Fredriksen for helping me out in laboratory. I will always be grateful to all other members of the group.

I want to thank all my other friends for supporting me and keeping me entertained during all these years. Without you time would be very different.

Alexey Bryukhovskiy

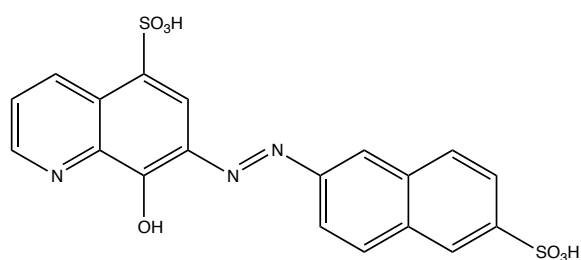
Abstract

c-Myb is a gene-specific transcription factor involved in the regulation of many target genes. As with other oncoproteins, an atypical function of the c-Myb may lead to uncontrolled cell growth and division that may cause many human cancers, where the c-Myb is over-expressed. Thus, it is observations like these that have established c-Myb as an attractive target for cancer therapy.

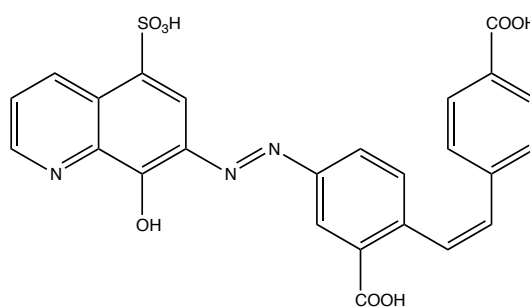
Halting expression of the c-Myb transcription factor was done by inhibition of the DNA-binding domain in c-Myb with small organic molecules. Compounds NSC87877 and NSC675460 (structures shown below) were previously identified as potential inhibitors of the DNA-binding domain in the c-Myb.

Five analogs of NSC87877 were synthesized. The abilities of the five compounds towards inhibition of the c-Myb transcription factor were tested both *in silico* and in biological assays.

In silico design of novel potential inhibitors based on the structure of NSC675460 was performed, resulting in 3 new lead compounds. Synthesis of compound NSC675460 was only partially completed.



NSC87877



NSC675460

Table of contents	Page
Abbreviations	5
Chapter 1. Introduction	7
1. The aim of the project	7
1.1 Introduction to the Myb-family of transcription factors	8
1.1.1 Transcription factors	8
1.1.2 The Myb family	9
1.1.3 The c-Myb transcription factor	9
1.1.4 The role of c-Myb in oncogenesis	10
1.1.5 Inhibition of transcription factors with small organic molecules	11
Chapter 2. Design and analysis of novel pharmaceutical agents	12
2.1 Background information	12
2.2 Structure-based drug design	12
2.3 Interactions between drug and target site	13
2.4 Software used in <i>in silico</i> design and analysis of molecules with drug-like properties	14
Chapter 3. AutoDock and AutoDock Tools	15
3.1 AutoDock	15
3.2 Force field function	15
3.3 Genetic algorithm	16
3.4 Grid maps and approximations	18
3.5 Essential concepts to note during analysis of docking results	18
3.6 Representative procedure for <i>in silico</i> analysis of potential drug compounds	19
Chapter 4. Overview and pharmacological analysis of NSC87877 and 675460	28
4.1 NSC87877	28
4.2 Metabolism of drugs	29
4.3 Pharmacological analysis of NSC87877	29
4.4 Pharmacological analysis of NSC675460	31
4.5 Lipinski's rule of 5	32
Chapter 5. Docking studies of NSC87877 and its analogues	33
5.1 Docking studies of NSC87877 and its analogues against the R2 region of DBD in c-	

Myb	33
5.2 Docking studies of NSC87877 and its analogues against the R3 region of DBD in c-Myb	36
Chapter 6. Computational studies of NSC675460 and its analogues	39
6.1 Docking of NSC675460 against the R2 and the R3 region in the DBD of the c-Myb	39
6.2 Bioisosteric approach in discovery of novel potential drugs	41
6.3 Bioisosteric replacement of the functional groups in NSC675460	42
6.4 Tetrazoles as bioisosteres of carboxylic acids	43
6.5 Docking of NSC675460 analogues against the R2 region of DBD in c-Myb	45
6.6 Docking of NSC675460 analogues against the R3 region of DBD in c-Myb	46
6.7 Re-docking	47
Chapter 7. <i>In silico</i> analysis of specificity of the compounds towards the DBD in c-Myb	51
7.1 Docking analysis of selected compounds against zinc finger 2 and 3 in the DBD of Sp1 protein	51
Chapter 8. Biological testing of the compounds AB74 – 78, AB81	53
8.1 EMSA-analysis	53
Chapter 9. Synthesis of compounds AB74, 75, 76, 77, 78 and 81	56
9.1 Theoretical aspects behind azo-coupling reaction	56
9.2 Synthesis, results and discussion	59
Chapter 10. Synthesis of the compound NSC675460	61
10.1 Introduction	61
10.2 Overview of NSC675460	61
10.3 Retrosynthetic analysis	61
10.4 Syntheses, results and discussion	63
10.4.1 Corey-Fuchs reaction	63
10.4.2 The Sonogashira coupling reaction	65
10.4.3 The Sonogashira coupling of the methyl 4-iodobenzoate with the ethynyltrimethylsilane	67
10.4.4 Removal of the TMS group	69
10.4.5 Preparation of the methyl 2-iodo-5-nitrobenzoate	70

10.4.6 The synthesis of methyl 2-(4-(methoxycarbonyl)styryl)-5-nitrobenzoate	72
10.4.7 The reduction of the nitro group and the hydrolysis of the ester groups	73
10.4.8 Introduction to the selective <i>cis</i> -reduction of the triple bond	75
10.4.9 Reduction of the triple bond using the Pd(OAc) ₂ /KOH/DMF–system	77
10.4.10 Reduction of the triple bond using the Pd(OAc) ₂ /PPh ₃ /NaOMe/THF–system	79
10.4.11 Synthesis of the <i>cis</i> -double bond using an addition/elimination approach	80
10.4.12 Synthesis of the <i>cis</i> -double bond using a hydroboration/protonolysis approach	80
10.4.13 Synthesis of the <i>cis</i> -double bond using a hydrosilylation/protodesilylation approach	82
10.4.14 Synthesis of the <i>cis</i> -double bond by hydrogenation using Lindlar catalyst system	86
Chapter 11. Conclusion	88
Chapter 12. Experimental	89
Attachments	110
Attachment 1. NMR data	110
(<i>E</i>)-8-hydroxy-7-((6-sulfonaphthalen-2-yl)diazenyl)quinoline-5-sulfonic acid	110
(<i>E</i>)-8-hydroxy-7-((5-sulfonaphthalen-2-yl)diazenyl)quinoline-5-sulfonic acid	114
(<i>E</i>)-8-hydroxy-7-((6-sulfonaphthalen-1-yl)diazenyl)quinoline-5-sulfonic acid	116
(<i>E</i>)-8-hydroxy-7-((5-sulfonaphthalen-1-yl)diazenyl)quinoline-5-sulfonic acid	119
(<i>E</i>)-8-hydroxy-7-((4-sulfonaphthalen-1-yl)diazenyl)quinoline-5-sulfonic acid	122
(<i>E</i>)-8-hydroxy-7-((7-sulfonaphthalen-1-yl)diazenyl)quinoline-5-sulfonic acid	125
Methyl 4-((trimethylsilyl)ethynyl)benzoate	128
Methyl 4-ethynylbenzoate	129
Methyl 2-iodo-5-nitrobenzoate	131
Methyl 2-((4-(methoxycarbonyl)phenyl)ethynyl)-5-nitrobenzoate	133
Methyl 5-amino-2-((4-(methoxycarbonyl)phenyl)ethynyl)benzoate	135
5-Amino-2-((4-carboxyphenyl)ethynyl)benzoic acid	137
(<i>E</i>)-methyl 2-(2-(4-(methoxycarbonyl)phenyl)-2-(triethylsilyl)vinyl)-5-nitrobenzoate	138
(<i>Z</i>)-methyl 5-amino-2-(4-(methoxycarbonyl)styryl)benzoate	139
Attachment 2. Miscellaneous NMR spectra	143
Attachment 3. Complete docking tables	151

Table 11	151
Table 12	154
Table 13	158
Table 14	165
Table 15	173
Attachment 4. AutoDock visualizations	175
Docking of AB81, 81-1 – 81-5 against the R2 region of the DBD in c-Myb	175
Docking of AB81, 81-1 – 81-5 against the R3 region of the DBD in c-Myb	176
Docking of AB74, 74-1 – 74-5 against the R2 region of the DBD in c-Myb	177
Docking of AB74, 74-1 – 74-5 against the R3 region of the DBD in c-Myb	178
Docking of AB75, 75-1 – 75-5 against the R2 region of the DBD in c-Myb	179
Docking of AB75, 75-1 – 75-5 against the R3 region of the DBD in c-Myb	180
Docking of AB76, 76-1 – 76-5 against the R2 region of the DBD in c-Myb	181
Docking of AB76, 76-1 – 76-5 against the R3 region of the DBD in c-Myb	182
Docking of AB77, 77-1 – 77-5 against the R2 region of the DBD in c-Myb	183
Docking of AB77, 77-1 – 77-5 against the R3 region of the DBD in c-Myb	184
Docking of AB78, 78-1 – 78-5 against the R2 region of the DBD in c-Myb	185
Docking of AB78, 78-1 – 78-5 against the R3 region of the DBD in c-Myb	186
Re-docking of ABAD45 in the R2 region in the DBD of c-Myb	187
Re-docking of ABAD45 in the R3 region in the DBD of c-Myb	188
Re-docking of ABAD53 in the R2 region in the DBD of c-Myb	189
Re-docking of ABAD53 in the R3 region in the DBD of c-Myb	190
Attachment 5. References	191

Abbreviations

ADMET	Absorption, Distribution, Metabolism, Excretion, and Toxicity
Ala	Alanine
Asn	Asparagine
Arg	Arginine
COSY	Correlated spectroscopy
d	Doublet
DBD	DNA-binding domain
dd	Double doublet
dt	Double triplet
DEPT	Distortionless enhancement by polarisation transfer
DMF	Dimethyl formamide
CH ₂ Cl ₂	Dichloromethane
CPU	Central processing unit
EI	Electron ionization
EMSA	Electrophoretic mobility shift assay
ESI	Electrospray ionosation
EtOH	Ethanol
Eq.	Equivalent
EtOAc	Ethyl acetate
Glu	Glutamic acid
Gln	Glutamine
Gly	Glycine
GUI	Graphical user interface
His	Histidine
HMQC	Heteronuclear Multiple-Quantum Correlation
HPLC	High performance liquid chromatography
IC	Inhibition constant
J	Coupling constant (Hz)
K _i	Inhibition constant
Lys	Lysine
m	Multiplet
MeOH	Methanol
MS	Mass spectrometry
<i>m/z</i>	Mass-to-charge ratio

μM	Micro-molar (concentration)
nM	Nano-molar (concentration)
NMR	Nuclear magnetic resonance spectrometry
OS	Operating system
PDB	Protein data bank
ppm	Part per million
q	Quartet
R_f	Retention factor
s	Singlet
Ser	Serine
td	Triple doublet
THF	Tetrahydrofuran
TLC	Thin Layer Chromotography
TMS	Trimethylsilane
TMSA	Ethynyltrimethylsilane
Trp	Tryptophan
ZF	Zinc finger

Chapter 1

Introduction

1. The aim of the project

- To investigate the nature of the transcription factor c-Myb in order to understand its role in many human cancers.
- To inhibit the expression of the transcription factor c-Myb by blocking the DNA-binding domain in c-Myb with small organic molecules.
- To discover novel inhibitors of c-Myb and evaluate its activity against c-Myb through the use of the specialized software.
- To investigate the possible metabolic routes of the newly discovered organic compounds.
- Use the bioisosteric approach to enhance the robustness of the compounds towards the metabolic processes in human body and biological activity of the molecule of interest.
- To investigate the selectivity of the newly discovered inhibitors towards the target molecule.
- To synthesize the newly discovered inhibitors and send them for testing in biological assays at the MYB group, the Department of Molecular Biosciences, University of Oslo.
- To compare results from the biological testing with results from computational studies. Evaluate how precise the software can predict the inhibition activity of the molecules of interest against the transcription factor c-Myb.

1.1 Introduction to the Myb family of transcription factors

The material in this chapter will present general information about the transcription machinery in eukaryotic cells, an overview of the Myb family of transcription factors, the proto-oncoprotein c-Myb, its biochemistry and its role in oncogenesis. These topics are necessary for understanding the strategy of inhibition of transcription factors with small organic molecules.

1.1.1 Transcription factors

Transcription is a process where a DNA-sequence is copied into a complementary RNA-sequence.¹ Enzymes that perform transcription of a DNA-segment are called RNA-polymerases. In eukaryotic organisms, there are three types of RNA-polymerases – I, II and III. RNA-polymerase II transcribes most genes, including those coding for proteins. A simplified process of a transcription can be viewed as follows: unwinding of a DNA-double helix, reading of a DNA segment and then catalysis of the formation of the phosphodiester bonds that link the nucleotides together to form a linear RNA chain. Initiation of transcription requires a vast array of general and specific transcription factors that bind to the promoters used by RNA-polymerase II, thereby making a sort of a platform for RNA-polymerase II.¹ The overall process of initiation of transcription can be illustrated in this manner (figure 1).

Gene-specific transcription factors are gene-regulating proteins that recognize DNA-sequences that code for particular proteins. They operate as a part of a bigger complex, where different co-factors interact with transcription factors thereby allowing better control of gene expression. Specific transcription factors have different specialized domains that perform specific tasks during transcription. A DNA-binding domain is responsible for recognition of specific site(s) on DNA and binding to a particular region. Another typical domain is one that mediates proteins activity.¹

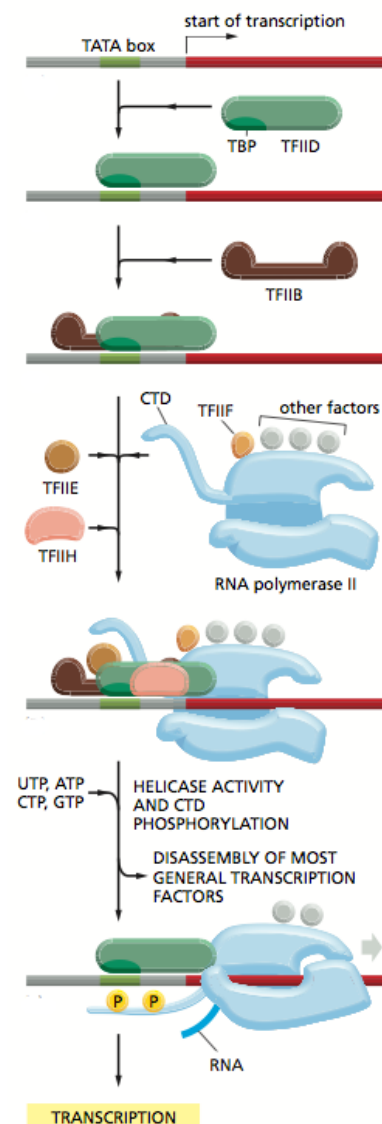


Fig. 1. Schematic representation of the transcription machinery. Fig. is taken from ref. 1.

1.1.2 The Myb family

The Myb-family is a family of transcription factors that are commonly found in eukaryotic organisms, including animals, fungi and plants.² The oncoprotein c-Myb is the best characterized member of the family that also consists of v-Myb, A-Myb and B-Myb. The main feature of the members of the Myb-family is the presence of so called “Myb domains” that are highly conserved regions located inside protein.² The Myb family was first discovered in its oncogenic form – v-Myb – in the AMV and the E26 retroviruses that cause monoclastic leukemia (AMV) and erythroblastic and myeloid leukemia (E26) in chicken.²

1.1.3 The c-Myb transcription factor

c-Myb is a gene-specific transcription factor with a helix-turn-helix (HTH) DNA-binding domain.³ The protein weights 75 kD and contains 640 amino acids in three different domains: The N-terminal DNA-binding domain (DBD), the central transactivational domain (TAD) and the C-terminal regulatory domain (CRD) (figure 2). These domains are highly conserved between members of the Myb-family and contain sites for interaction with DNA and other proteins. In the scope of this project, only the DNA-binding domain will be thoroughly discussed.

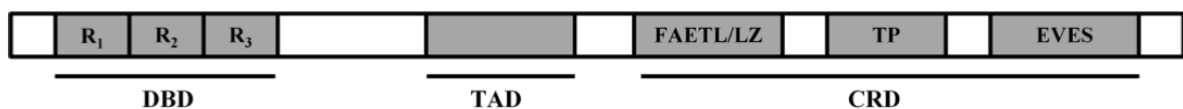


Fig. 2 Schematic representation of c-Myb's functional domains. The DNA-binding domain is located in N-terminus, in the central part lays transactivational domain and CRD is located in C-terminus.

The DBD in c-Myb has a HTH-like motif and consists of three tandem repeats that are commonly referred to as R1, R2 and R3 starting from the N-terminal.⁴ Each repeat contains 51 – 52 amino acids including three tryptophan residues that are located after every 18-19 amino acids. The Trp residues are crucial for DBD's ability to interact with the DNA since they maintain HTH-like structure that is critical for the correct positioning of the DNA sequence. In addition, acidic and hydrophobic regions are also found in the domain.

The R1, R2 and R3 regions consist of three α -helices that are folded and create a complex structure (figure 3).^{3,5} It has been shown that only the R2 and R3 repeats bind directly to the DNA, while the R1 stabilize the protein/DNA structure through electrostatic interactions (figure 3). When DNA comes in contact with the R2 and R3, hydrogen bonds between the Asn183, Asn179 and Lys182 residues from the R3 region are formed with the following bases from DNA: A18, A19 and G8. In the R2 region

hydrogen bonds are created between Glu132 and Lys128 and bases G22 and C20 respectively (figure 4).³

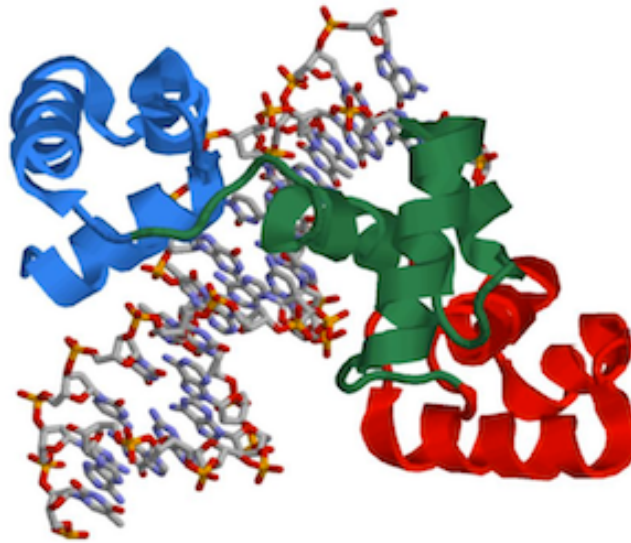


Fig. 3 Complex between R1, R2, R3 and DNA. R2 is green, R3 is blue and R1 is red. As it can be seen, R1 does not make direct contact with the DNA chain. Fig. is taken from ref. 6.

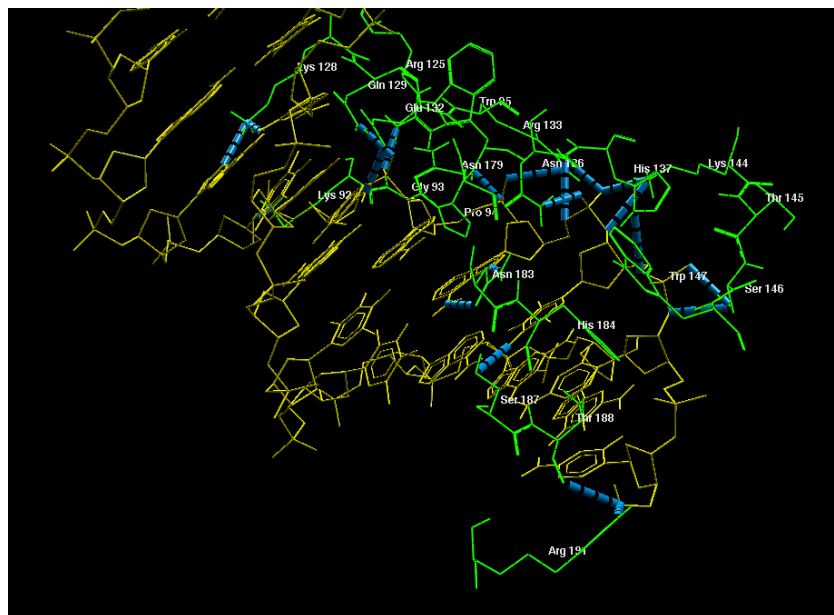


Fig. 4 Hydrogen bonds (blue) between DNA (yellow) and the residues from the R2R3 (green). Image was generated with MMV.

1.1.4 The role of the c-Myb in oncogenesis

Bone marrow, large intestine and mammary glands are primary targets of the c-Mybs oncogenic activity in humans.² The oncogenic activity of c-Myb can be activated by over-expression or inappropriate

expression, structural alteration and/or genomic rearrangements.⁷ The genomic locus of the c-Myb gene is often rearranged in human acute myelogenous leukemias, melanomas and carcinomas. This rearrangement of the locus results in an amplification of the expression of the c-Myb gene. Increased expression of c-Myb has also been detected in other cancers, such as breast and colon cancer. Genomic rearrangements in the c-Myb gene have also been identified in T-cell acute lymphoblastic leukemia, where duplication of the c-Myb gene and translocation of the c-Myb locus led to increased proliferation and blocking of differentiation.

c-Myb is crucial in several processes required for carcinogenesis.⁷ It has been shown that high levels of c-Myb expression are necessary for survival and expansion of several types of human cancers. Leukemia cells are also more sensitive towards inhibition of c-Myb than normal hematopoietic cells. All these features make c-Myb an attractive target for the development of potential medicinal agents against human cancers.

1.1.5 Inhibition of transcription factors with small organic molecules

The general transcription machinery is used by all transcriptional factors, meaning that inhibition of some part of this machinery will have an affect on the whole system.^{8,9} In order to modulate the activity of a chosen transcription factor, a small organic molecule must bind itself to an area on a surface of transcription factor that is critical for its normal functioning. The mode of action of small molecules as inhibitors can be divided into several classes:⁹

- Small molecules that bind to the DNA binding domain of transcription factor(s).
- Small molecules that bind to the DNA sequence targeted by transcription factor(s).
- Small molecules that prevent dimerization of the transcription factor(s).
- Small molecules that prevent contact between co-activators and transcription factor(s) by binding to the protein-protein interaction domain.

c-Myb acts as monomer, making prevention of dimerization pointless in this case. Martinez et al. chose inhibition of the DNA binding domain in c-Myb as the approach of his choice and this approach was followed in this study also.⁶⁸ The general idea was to find small organic molecules that can make a tight complex with residues involved in binding of the DNA strand, thus preventing contact between the DNA binding domain in the c-Myb and the DNA.

Chapter 2

Design and analysis of novel pharmaceutical agents

In this chapter, some background information about drug design is given. It includes an overview of major guidelines that are followed during the design of molecules with drug-like properties using *in silico* tools.

2.1 Background information

A drug is usually a small molecule that binds to, interacts with and affects the activity of a target molecule. Drugs are commonly used to identify, prevent or treat a disease. Pharmaceutical companies spend billions of dollars on discovery of novel drugs – a process that can easily take up to 10 years.¹¹ Discovery of new drugs has historically been a trial and error process where hundreds of chemical compounds have been screened in biological assays.¹¹ The analysis of a drug candidate can be separated into several parts: examination of biological activities and ADMET parameters of a drug. The discovery of a possible drug is a multistep process that can roughly be divided into three stages: target identification, lead discovery and clinical trials.¹¹

The drug target is a macromolecule that is critical for a disease to function. Target molecule may be an enzyme, receptor or a transcriptional factor.¹⁰ Detailed knowledge about the target comes from biological and genetic investigations, X-ray and NMR studies. Lead discovery has been a very expensive stage of drug discovery process until 1980's since it required very large amount of experimental work and time. In the 1980's, computer aided drug design began to expand and grow due to progress made in the field of chemo- and bioinformatics as well as computer power. Today, large pharmaceutical companies have chemoinformatic departments, which enable them to screen millions of potential organic compounds.¹¹ Chemoinformatic tools saves both time and money, it also helps researchers to avoid experimental dead ends that are not trivial to predict. There are several approaches towards discovery of new potential drug candidates using computer aided drug design. Structure based drug design (SBDD) is one of the approaches and it also the one that was used in this project.

2.2 Structure-based drug design

Structure based drug design relies on a fact that the structure of the target molecule is known.¹⁰ A potential drug candidate can then be discovered by screening of large databases of organic compounds. The initial process of finding a new drug candidate consists of two main parts: docking and scoring.

Docking is the generation of many different orientations of a drug candidate in and around the binding site of the target molecule. Using software language, it is a genetic algorithm that performs the docking. During docking, the software makes many spatial transformations of drug and receptor, trying to find most energetically favourable orientations. Scoring is the prediction of how well a particular orientation of a drug will be bonded to the target. Since inhibitor potency is directly related to the degree of target molecule–drug binding, it means that scoring is one of the main tasks in accurate prediction of the target-inhibitor affinities. The conformation with the lowest energy is usually the one that is predicted as the binding mode.

2.3 Interactions between drug and target site

The material in this section (section 2.3) is based on previously published works.^{1, 11, 73}

When a molecule binds to a target substrate, vast arrays of interactions arise.¹¹ Water molecules get expelled upon binding, making it possible for functional groups to interact with each other. The binding of a drug to a target and how strong it will be depends solely on what type of bonds are being formed. The four most common types of bonds are covalent, ionic, hydrogen and London-forces (figure 5).

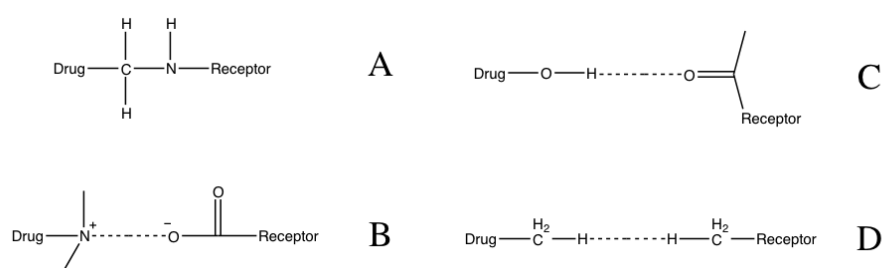


Fig. 5. Schematic representation of the interactions. A – covalent bond, B – ionic bond, C – hydrogen bond, D – London-forces

The strongest bond that can be found in drug-target interactions is covalent bond, where an atom from drug shares a pair of electrons with an atom from target molecule. Covalent bond is very strong (typically 50 – 100 kcal/mol) and often results in irreversible binding of drug to receptor.

The next-strongest bond that can be formed between drug and target molecule is ionic bond. The strength of the ionic bond lies between 5 and 10 kcal/mol and its strength decreases proportionally to the square of the distance of two atoms that are involved in bonding. Ionic interactions increase significantly when drug differs closer to target. Ionic bond does not result in irreversible binding – a drug molecule can at some point diffuse away from the target.

Hydrogen bonds are formed when a hydrogen atom is shared between two atoms, an electronegative donor and a basic acceptor atom. The most important hydrogen bonds are those involving the oxygen and nitrogen atoms of the carboxyl, hydroxyl, carbonyl, amino, imino and amido groups. A single hydrogen bond has on average strength of 2 – 5 kcal/mol, which is usually not sufficient enough to stabilize a drug-target complex on its own. More commonly a typical interaction between a drug and a target includes several H-bonds, that all confer to the stability of the drug-receptor complex.

The weakest interactions that can be found in drug-target complex are London and hydrophobic interactions. London-forces appear from fluctuations in electron density around atoms. Unsymmetrical / disturbed electron densities around atom(s) causes a similar electron density to occur on neighboring atom(s) resulting in an attraction. These forces become significant as distance between clouds of atoms decreases. Hydrophobic interactions are forces that can be quite significant in aqueous solutions – they are formed when water molecules are expelled from the surface between drug and target molecules upon binding. Hydrophobic interactions are entropy driven – expelling water results in increase in entropy and decrease in energy.

2.4 Software used in *in silico* design and analysis of molecules with drug-like properties

There are many programs that are designed for the screening and evaluation of small molecules as potential drug candidates. The main difference between the programs comes from what type of function is being used for the scoring of ligands. Various scoring functions offer different accuracies in the scoring process, meaning that some programs provide theoretical results that are close to experimental values, while other software programs do not. In the past 20 years, numerous scoring functions have been developed and they can be divided into three major groups: force field,¹² empirical¹³ and knowledge-based scoring functions.¹⁴ The choice of the software is regrettably not always based on their scoring accuracy. Factors like economic cost and user-friendliness are major concerns. Many programs do not have a GUI, so users very often spend more time learning codes and programming languages than actually performing scientific research. The compatibility of the software with CPUs, graphic cards, OSs etc is another major issue.

Computational part of this project was based on a previous work done by Martinez et al.⁶⁸ The software for docking studies that was used both in this and the previous project is AutoDock 4.2.¹⁵ Reasons for keeping AutoDock as our software of choice, is that it is a free software package that can be installed both on Mac OS X, Linux and Windows platforms and thus results from our and previous studies⁶⁸ could to some extent be compared.

AutoDock and AutoDock Tools

In this chapter, a presentation of the AutoDock package will be given. General steps in docking process will be reviewed. A short introduction on how to use the software will also be given.

3.1 AutoDock

AutoDock is a software package consisting of AutoDock (AD) and AutoDock Tools (ADT). The latest stable versions of AD and ADT are 4.2 and 1.5.4. AutoDock is a program that predicts the optimal bound conformation of a ligand-protein complex. AutoDock uses a semi-empirical free energy force field for scoring combined with a suitable genetic algorithm for generation of potential orientations of a drug molecule. A number of approximations are used during the process of finding and evaluating conformations to minimize computational time.

3.2 Force field function

The semi-empirical free energy force field estimates energy of binding of two (or more) molecules in a water environment by using pair-wise terms to evaluate the interaction between two molecules and an empirical method to estimate contribution of the surrounding water.¹⁶ The goal of this force field is to capture enthalpic and entropic contributions to the drug-target complex in a limited number of easily evaluated terms. The free energy of binding is the difference between the energy of ligand and the protein in an unbound state and the energy of ligand-protein complex. This difference is broken further in to two pieces/steps: evaluation of intramolecular energies of the transition from the unbound state to the bound one for each of the molecules separately, and then evaluate the intermolecular energetics of bringing two molecules together into a bound complex (figure 6).

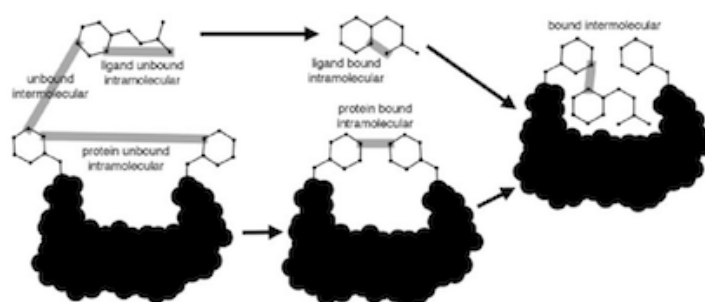


Fig. 6. Schematic representation of scoring process in AutoDock 4. Fig. is taken from ref. 16.

The estimation process of the free energy of bonding of one particular conformation includes six pairwise evaluations (V) and estimation of the lost conformational entropy upon binding (ΔS_{conf}), L stands for ligand and P for protein (equation 1):¹⁶

$$\Delta G = (V_{\text{bound}}^{L-L} - V_{\text{unbound}}^{L-L}) + (V_{\text{bound}}^{P-P} - V_{\text{unbound}}^{P-P}) \\ + (V_{\text{bound}}^{P-L} - V_{\text{unbound}}^{P-L} + \Delta S_{\text{conf}})$$

Equation 1

Every evaluation V is then composed of evaluations for dispersion/repulsion, hydrogen bonding, electrostatics and desolvation (equation 2):¹⁶

$$V = W_{vdw} \sum_{ij} \left(\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} \right) + W_{\text{hbound}} \sum_{ij} E(t) \left(\frac{C_{ij}}{r_{ij}^{12}} - \frac{D_{ij}}{r_{ij}^{10}} \right) \\ + W_{\text{elec}} \sum_{ij} \frac{q_i q_j}{\epsilon(r_{ij}) r_{ij}} + W_{\text{sol}} \sum_{ij} (S_i V_j + S_j V_i) e^{(-r_{ij}^2/2\sigma^2)}$$

Equation 2

Constants – W etc. – have been calibrated by using a large dataset of known ligand-target complexes, where the free energy of binding and other parameters have been obtained experimentally.

3.3 Genetic algorithm

The second major part of AutoDock software is the genetic algorithm (GA) used for docking. As it was mentioned before, docking is a process of finding new conformations of protein and target molecules. Exploration of possible conformations through the use of genetic algorithm is inspired by natural genetics and biological evolution. As implied by the name, the problem of discovery of new energetically favorable orientations of ligand and target molecule must be encoded using “genetic language” – as a genome or chromosome. Genomes or chromosomes undergo common evolutionary “events” - mutations, crossovers etc – slightly improving its “properties” – finding conformation with lowest energy. Lamarckian genetic algorithm proved to be one of the best genetic algorithms for use in docking experiments with large conformational space. The material in this section is based on previously published material.¹⁷

The first step in a docking process is to define a particular arrangement of ligand and protein using state variables: translation of the ligand (x, y and z coordinates in space), orientation and conformation of ligand with respect to the target molecule. Each state variable of ligand corresponds to a “gene”, so all state variables form ligand’s “genotype”, while atomic coordinates of ligand correspond to the

“phenotype”. When the process of docking starts, each genotype is evaluated by a scoring function that gives total interaction energy of ligand and target molecule – using language of GA this process is called fitness. Another important term used in GA is “chromosome” – a set of all parameters that describe both ligand and target molecule – its genotype and phenotype.

In a docking experiment, the “chromosome” of a compound is thus made of several genes: three Cartesian coordinates for the ligand translation, four variables defining the ligand orientation and one value for each torsion of the ligand. Once the genes have been defined, the genetic algorithm starts by creating a population of individuals where the number of individuals in the population is defined by the user. Process of crossover and mutation are applied to create new individuals in a population by changing one or more of the state variables – three translation genes for x, y and z is given a random value between the minimum and maximum x, y and z extents of the grid map, the four genes describing the orientation are given a random quaternion, consisting of a random unit vector and a random rotation angle between -180° and +180°. Then a process of “fitness” of every individual is applied using the scoring function. Like in biology, not all “individuals” survives – only the best go to the next step in the process of docking/scoring in accordance with (equation 3):

$$n_o = \frac{f_w - f_i}{f_w - \langle f \rangle} \quad f_w \neq \langle f \rangle$$

Equation 3

In equation 3 n_o is the integer number of offspring to be allocated to the individual; f_i is the fitness of the individual, that is, the energy score of the individual; f_w is the fitness of the worst individual (highest energy) in the last N generations (i.e., N is a user-definable parameter, typically 10) and $\langle f \rangle$ is the mean fitness of the population. Since f_w will always have a higher energy than both f_i and $\langle f \rangle$, except for when $f_w = f_i$. For individuals who have a fitness lower than the mean $f_i < \langle f \rangle$, the numerator in this equation, $f_w - f_i$, will always be greater than the denominator $f_w - \langle f \rangle$, and therefore such individuals will be allowed at least one offspring, and thus will be able to reproduce. In this way the more “fit” individuals will be allowed to carry their “genes” into the next generation, whilst the least “fit” will eventually become “extinct”. Finally if $f_w = \langle f \rangle$, the population is assumed to have converged and the docking is terminated.

When individuals for the next generation are selected, the process of crossover, mutation and scoring starts all over again. This process is repeated for a number of generations that is user-defined. Sometimes random mutations and crossovers may drive the search for a global minimum in a wrong direction. Lamarckian GA deals with this issue by replacing state variables in the genotype with those obtained in the last docking run that corresponds to the lowest energy.

3.4 Grid maps and approximations

A fast evaluation of interactions between ligand and target molecule is achieved through the use of grid maps (figure 7).¹⁸ It means that affinity potentials are pre-calculated for each atom of a ligand. The target protein is set in a three dimensional grid and a probe atom is placed at each grid point. The energy of interaction is assigned to this point. An affinity map is calculated for every atom and every single grid point that is present in the potential drug candidate. A grid map of electrostatic potentials is pre-calculated too. Subsequently these maps are used as a look-up table in order to perform a rapid energy evaluation.

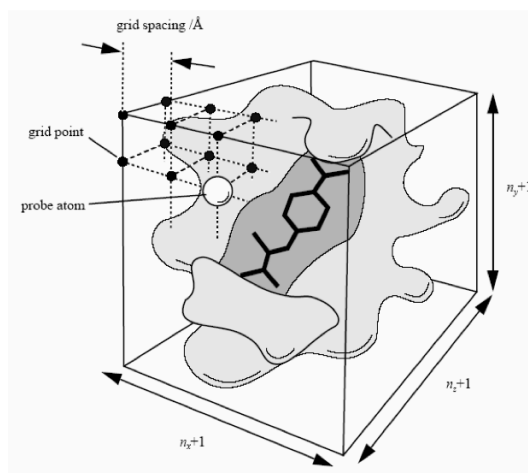


Fig. 7. Representation of the grid box concept. Fig. is taken from ref. 21.

In order to provide accurate results in a reasonable amount of time, a number of approximations must be included in the method of docking. Treating the target molecule as rigid is one of the most common simplifications. However, it is known that upon binding of a ligand to a target molecule some conformational changes occur, ranging from local side chain movements to large domain changes. The target molecule itself is not a rigid structure where all atoms and chains are fixed at some position. There is always some amount of flexibility and structural fluctuations. This type of motion can be visualised / quantified upon performing NMR analysis of a macromolecule – all 20 structures from 20 “runs” will slightly differ from each other in terms of bond lengths and angles between side chains. As it was mentioned above, AD’s “grid map” approach implicates that the target molecule is a rigid structure with zero degrees of freedom. One way to overcome this problem is to put specific side-chains outside of the grid box and treat them as flexible.

3.5 Essential concepts to note during analysis of docking results

Experimental affinities between a drug-like molecule and a target molecule usually differ from those obtained from docking software. So, if docking software cannot accurately predict experimental values,

it must at least rank the compounds to be screened in the correct order. Reproducibility of the binding mode and location is also important. AD has standard deviation of free energy prediction of about 2 - 3 kcal/mol.¹⁷ Unfortunately it is not sufficient precision to enable ranking of the compounds as the best. Often, best hits from the first round of docking have to be “re-docked” to ensure that the pose and affinity is not just a statistical coincidence.

One of the most important steps in docking analysis of potential drug candidates is the evaluation of the docking results. First of all, the software and the docking algorithm must be tested on a known system where the actual drug has subsequently been investigated by X-ray crystallography to ensure that the binding pose suggested by the software does not differ much from the experimental one. IC 50 values may also be compared. AD ranks docked compounds in order of increasing free energy of binding, which is a good criterion for a fast overview of results. Additionally, analyses of consistency of docked conformations by clustering of compounds based on RMSD (root mean square deviation) of the coordinates may provide additional proof. Conformations that are found many times in the same region of the target molecule are grouped into clusters. Clusters are then separated from each other according to their free energy of binding. So, when a drug-like molecule binds to a target molecule, it becomes less mobile and loses its entropy, which opposes binding forces between target and ligand.^{19, 20} It means that a cluster with a large quantity of molecules is more likely to induce favorable entropy changes in the system. If a cluster with a small amount of molecules has a lower free energy than a big cluster with higher free energy, then it is more probable that a large cluster will represent a more or less “true” picture of binding pose. Another important factor is the so-called ligand efficiency criteria.²² Large molecules tend to have better predicted affinity than smaller ones due to the fact that the binding affinity is somewhat proportional to the number of atoms in the ligand. Ligand efficiency calculates free energy of binding per non-hydrogen atom, giving a rough estimate of how “efficient” the molecule is. This being said, a large molecule with an affinity that is only 1 kcal/mol lower than another molecule that is twice as small, is not necessarily a better drug candidate.

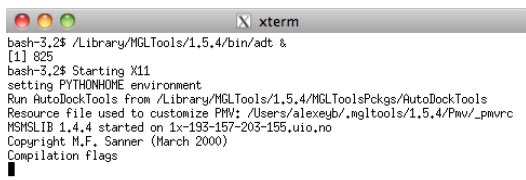
3.6 Representative procedure for *in silico* analysis of potential drug compounds

A single docking is a process where a single molecule is docked against a target molecule. From this point on, the word “docking” will be used about the whole process of evaluation of molecule with AD and ADT. To perform a single docking, one needs 4 files:

1. A file describing the ligand
2. A file describing the macromolecule
3. A file describing the search parameter file
4. A file describing the parameters of the search

Preparation of the macromolecule file

The structure of c-Myb DBD can be downloaded from the PDB's databank website <http://pdb.org>. The file name is 1GV2.pdb. Open ADT 1.5.4 either by clicking on an icon or by typing in terminal: `/Library/MGLTools/1.5.4/bin/adt &` and press enter (figure 8). PDB-files are not perfect – sometimes they include water molecules and other atoms that must be removed prior to use. Click on *File* → *Read molecule*, open 1GV2.pdb (figures 8, 9, 10).



```
bash-3.2$ /Library/MGLTools/1.5.4/bin/adt &
[1] 825
bash-3.2$ Starting X11
setting PYTHONHOME environment
Run AutoDockTools from /Library/MGLTools/1.5.4/MGLToolsPkgs/AutoDockTools
Resource file used to customize PMV: /Users/alexeyb/.mgltools/1.5.4/PMV/_pmwrc
MSHSLIB 1.4.4 started on 1x-195-157-203-195.uio.no
Copyright H.F. Sanner (March 2000)
Compilation flags
```

Fig. 8

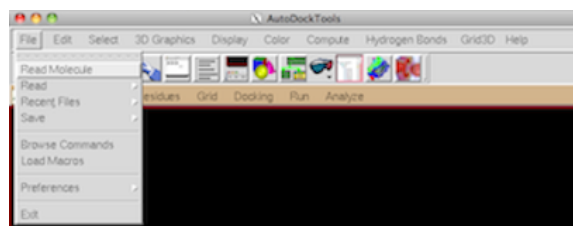


Fig. 9

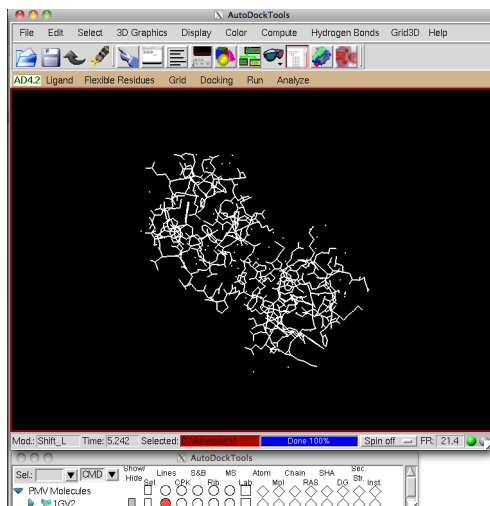


Fig. 10

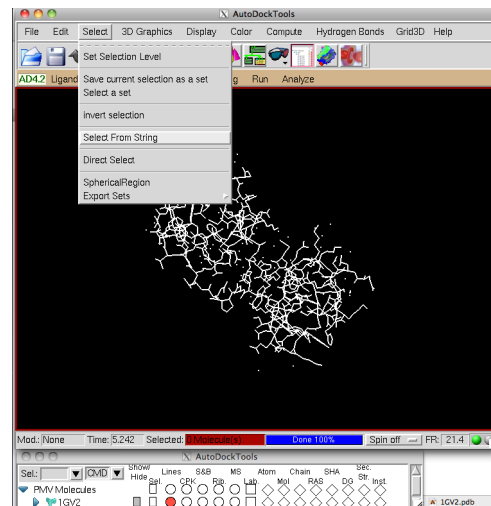


Fig. 11

In order to remove the Na^+ ions and water, click *Select* → *Select from string* (figure 11). Type *HOH** in the *Residue entry* and *** in the *Atom entry*. Click *Add*, this will select 60 atoms (figure 12).

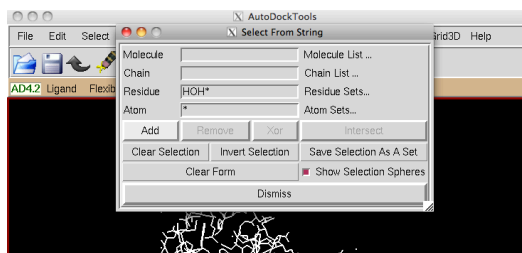


Fig. 12

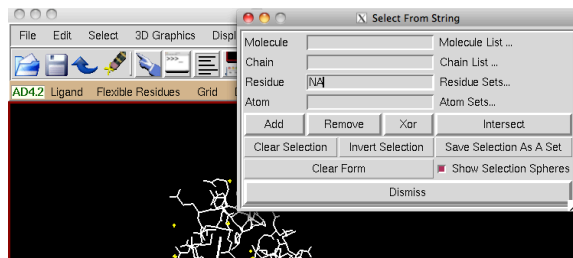


Fig. 13

Click on *clear form* and type *NA* in the *Atom entry*. Click *Add*. You will see *Selected: 62 residues*

(figure 13). Then click *Dismiss* in the *Select From a String* window and click *Edit* → *Delete* → *Delete Atom Set* (figure 14). You will get a warning message, click on *CONTINUE*. This will delete all water molecules and the 2 Na⁺ ions. Then click *Edit* → *Hydrogens* → *Add*, choose to add *All Hydrogens* using method *noBondOrder* with *yes* to renumbering. Click *OK* to add all hydrogens. Then click *File* → *Save* → *Write PDB*. Click *OK* to write the file. Then click *Grid* → *Macromolecule* → *Choose*, choose 1GV2 → *Select Molecule*. A new window will be opened where you will be asked to save your macromolecule file in the PDBQT-format.

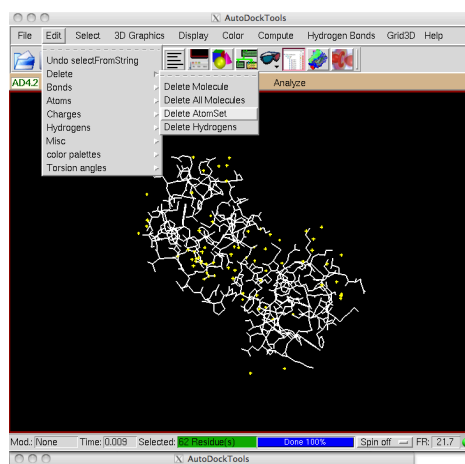


Fig. 14

If the molecule already has charges, the program will ask you if you want to preserve the charges instead of adding Gasteiger charges. Click *Yes*. Then click *Edit* → *Delete* → *Delete all molecules* and click *continue*

Preparation of the ligand file

Click *Ligand* → *Input* → *Open* and choose the .pdb file corresponding to the molecule you want to dock. Upon opening the ligand molecule, ADT automatically computes Gasteiger charges, merges non-polar hydrogens, lists how many rotational bonds molecule of your choice has, and detect aromatic carbons. It also colors atoms according to their type (carbon, oxygen etc) (figures 15, 16).

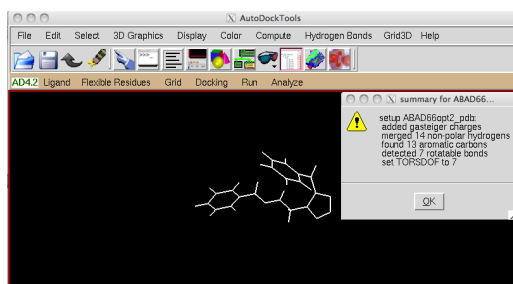


Fig. 15

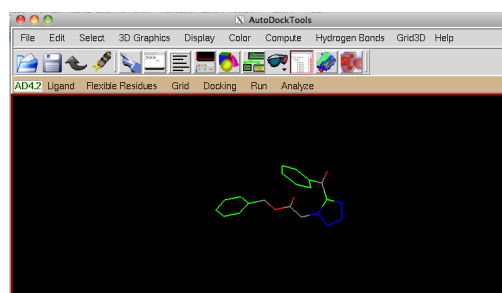


Fig. 16

Then the ligand must be saved as a PDQT-file. Click *Ligand* → *Output* → *Save as .pdbqt* (figure 17).

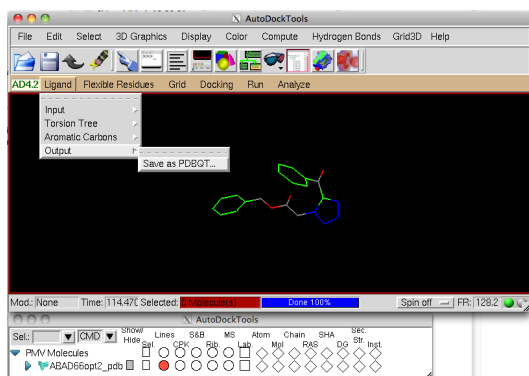


Fig. 17

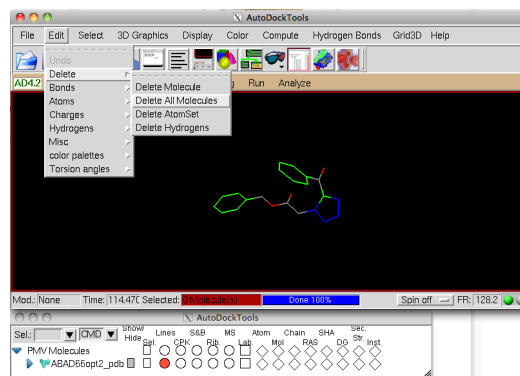


Fig. 18

Press *Edit* → *Delete* → *Delete all molecules* and click *Continue* (figure 18). Right now both the macromolecule and the ligand files are prepared. The next step is to prepare the grid parameter file and the parameters search file. Click *Grid* → *Macromolecule* → *Open* and find 1GV2.pdbqt, then click *Open* (figure 19).

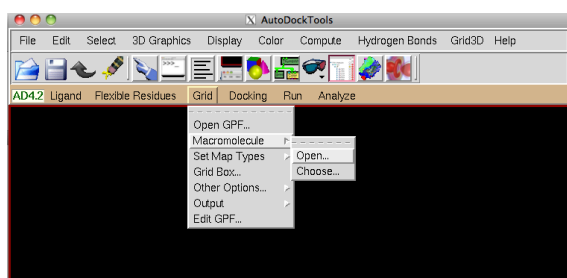


Fig. 19

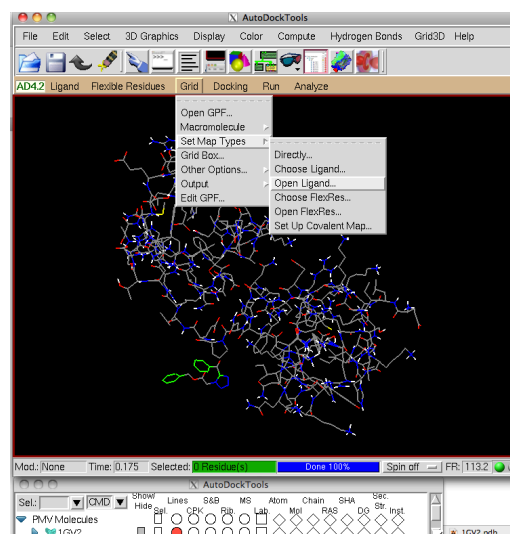


Fig. 20

You will get a warning message where you will be asked if you want to preserve the already added charges etc. Click *No*. Then Click *Grid* → *Set Map types* → *Open Ligand*, find ligand.pdbqt file and click *Open* (figure 20). If everything was done according to this procedure, you will see both the ligand and the macromolecule in the ADT window. Next step is to generate a grid affinity map for the R2 or R3 region. As it was pointed before, the residues that are important for DNA binding in the R2 region are Glu132 and Lys 128 and for the R3 region - Ser187, Asn183, Asn179, Lys183 and Asn186. The grid boxes must be centered on those residues and have a sufficient size so ligand molecules can be rotated in all directions. Lys 128 and other residues can be located by clicking on *Select* → *Select from string*,

in the *Residue* field type Lys128 and click *add*. Lys128 will then be highlighted with small yellow asterisks. Based on those criteria, the center and the size of the R2 grid box are x: 22.153, y: 13.581, z: 14.192; size: 50 x 40 x 60. The center and the size of the R3 grid box are: x: -0.503, y: 15.076, z: 7.389; size: 74 x 44 x 40. In order to understand those numbers and what they mean, we shall proceed to the “construction” of the grid box. In order to dock our drug like molecule against the R2 region, click *Grid* → *Grid Box...* and the *Grid Options Window* will come up together with a box (figure 21).

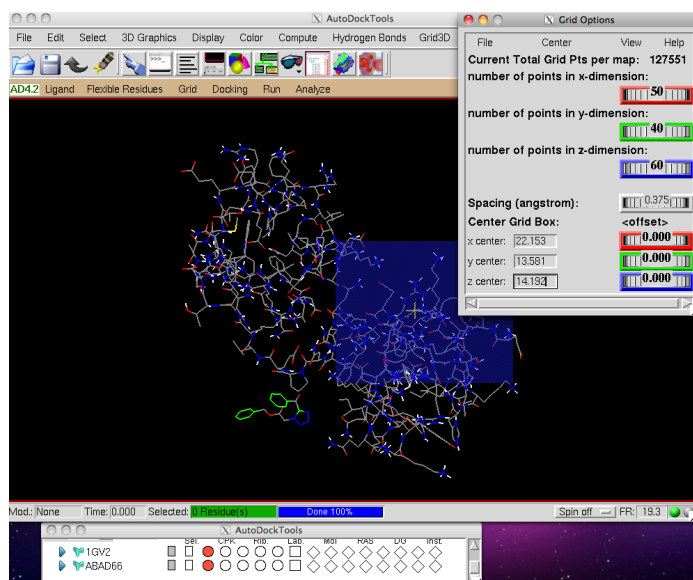


Fig. 21

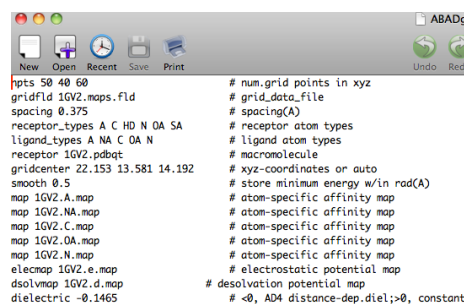


Fig. 22. GPF-file can be opened with any emacs editor.

There are several important things to notice here. By clicking on *Center* → *Pick an atom* and picking an atom that belongs to for e. g. Lys128, the grid box will center on the Lys128. The number of points in the x, y and z dimensions correspond to the size of the grid box. Typing the coordinates above in the center grid box fields can also center the grid box. Adjust number of points in the x, y and z direction to 50 (x), 40 (y) and 60(z) by moving the “wheels” (the ones with the number 40 typed on them). Then type 22.153, 13.581 and 14.192 in the *Center Grid Box* fields (figure 23). If everything was done correctly, then the *Grid Options Window* will look like it does in the figure 23. Then click (in the *Grid Options Window*) *File* → *Close Saving Current*. Then close the *Grid Options Window*. Then click *Grid* → *Output* → *Save gpf* and give the .GPF-file an appropriate name. A GPF-file is actually a text file (figure 22).

Preparation of the search parameter file

Click on *Docking* → *Macromolecule* → *Set rigid file name...* choose 1GV2.pdbqt. Click on *Docking* → *Ligand* → *Choose...*, choose “ligand.pdbqt” and click *Close* in the window that will appear. Click on *Docking* → *Search parameters* → *Genetic algorithm...*, this will open a window called *Genetic*

Algorithm Parameters. According to the literature one should make an output file with 100 GA runs, a value of 150 in population size and 10 million of energy evaluations.²³ To reduce the time of the calculations, use 5 million of energy evaluations. Click *Accept*. Click *Docking* → *Docking parameters*, this will open a window called *Set Docking parameters*, change translation step to 1.0. Click *Accept*. Then both windows will look like they do in figure 23. Click on *Docking* → *Output* → *Lamarckian GA* and give the DPF-file an appropriate name.

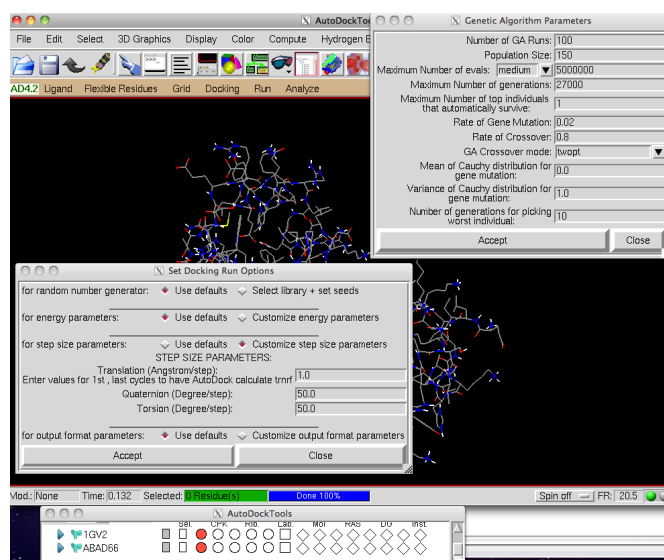


Fig. 23

Docking of the ligand against the DBD of the c-Myb

The actual docking process using the AutoDock software is done at the Titan cluster. The first thing to do is to log on at <http://bioportal.uio.no>, click on *Projects* and create one with an appropriate name (figure 24).



Fig. 24

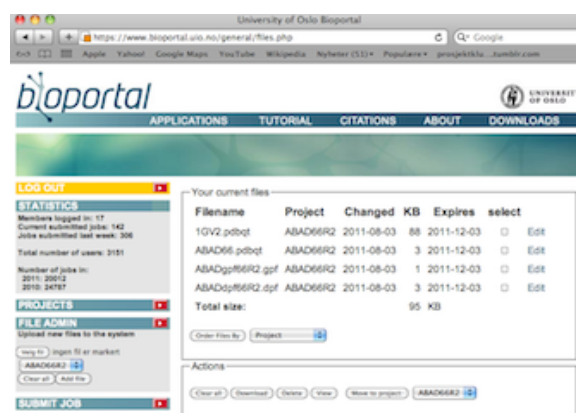


Fig. 25

Upload all the 4 files to the project in *File Admin* (figure 25). Click on *Submit job*, click on appropriate project in the *Choose project* field and mark your desired input files. In the *Resource* roll-menu click on *titan*. In *Application*, choose *autodock4@titan* and enter the name of the job in *Job name* (figure 26). Click on *Submit*. This will lead you to the screen depicted in figure 27.

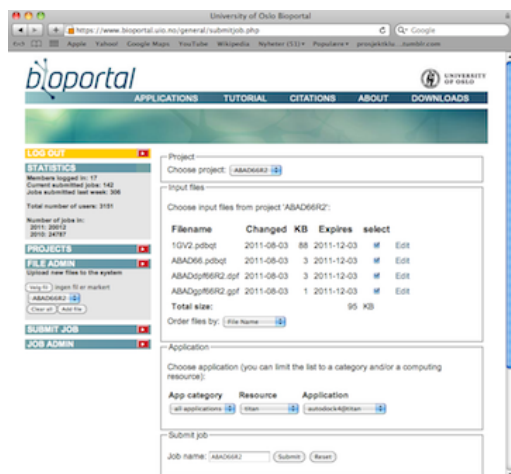


Fig. 26

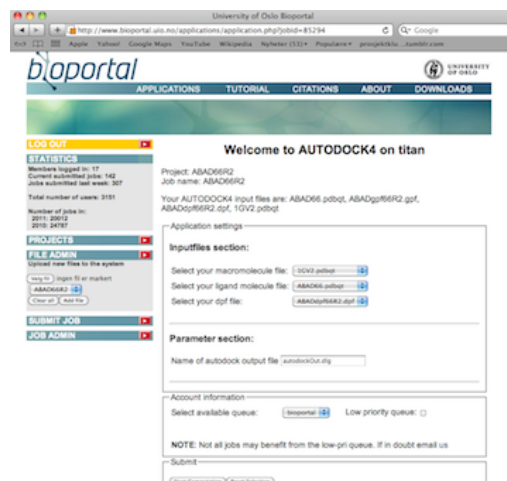


Fig. 27

Specify your macromolecule, the ligand and the DPF-file in the *Inputfiles* section, in the *Select available queue* select *biportal* and unmark *Low priority queue*. Then click on *Start computation*. Now the docking process / computation will start. It takes about 6 - 8 hours to run a docking of one compound in one region.

When the job is done, all the files associated with the docking run can be downloaded as a .zip file. Unzip the files to the same directory the input files are located. Then, in ADT, click on *Analyze* → *Dockings* → *Open...*, find the unzipped output files and open the file *autodockOut.dlg* and click on *OK* (figure 28).

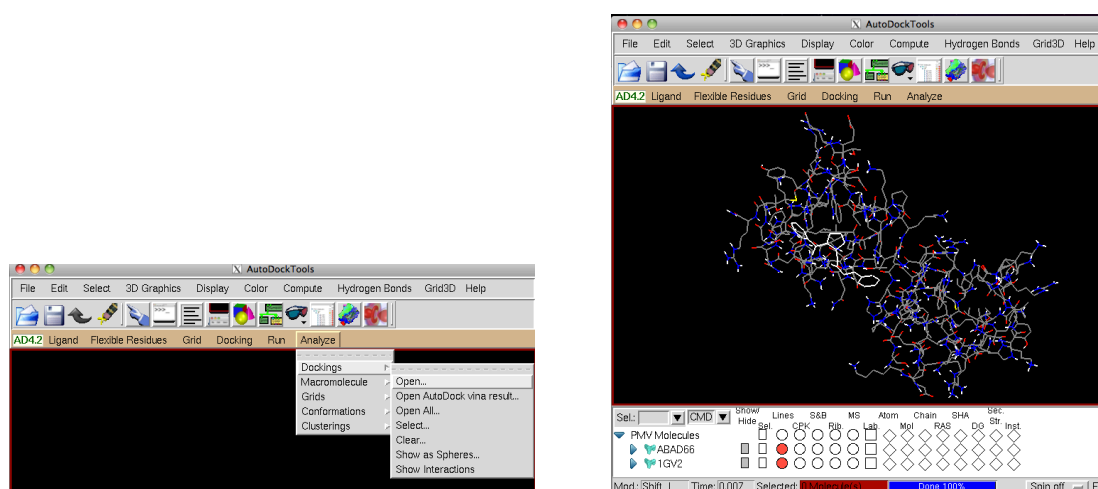


Fig. 28

Fig. 29

Then open the macromolecule file 1GV2.pdbqt – Click on *Analyze* → *Macromolecule* → *Open...*

If everything was done correctly, both the macromolecule and the ligand will appear in the ADT window. The next step is to load the most important thing – the docked conformations and affinities, Click on a circle under *MS* and on a square under *Atom* that corresponds to the macromolecule in order to visualize the molecular surfaces and colour them by atom type. Click on *Analyze* → *Conformations* → *Load*, this will bring up the *Autodock conformation chooser* window (figure 30).

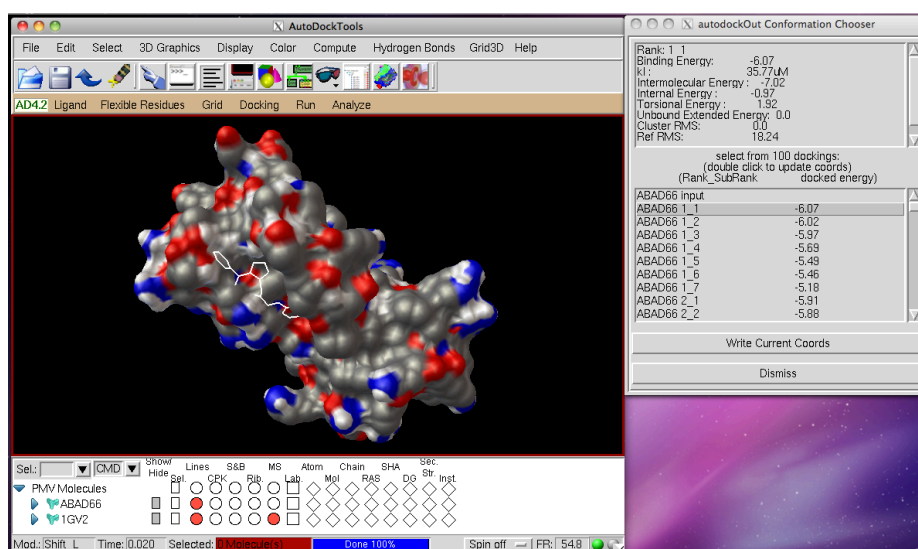



Fig. 30

By clicking on *ABAD66 1_1*, the conformation with the lowest binding energy will come up in the ADT window together with its corresponding inhibition constant. In number 1_1, the first digit corresponds to the number of the cluster based on the RMSD, the second digit corresponds to the molecule with the lowest energy – *ABAD66 2_2* means that it is cluster number 2 and this docked conformation has the second lowest energy. Clustering based on the RMSD values can also be viewed by clicking on *Analyze* → *Clusterings* → *Show*. Two windows will appear – one called *ABAD66: rms = 2.0 clustering* and another – *ABAD66*. Second window – *ABAD66* – is a conformational player that allows going through different conformations by clicking on the arrows. In order to view additional information about the different docked ligands, the following can be done.

Click on the  button and the *Set Play options* window will come up, in that window click on *Show Info* and *Build H-bonds*. Two more windows will come up, one called *Conformation X_X info* and another called *Hydrogen bonds*. In the first window, all information about energies, efficiency etc. can be found. In the *Hydrogen bonds* window the residues on the macromolecule that are involved in hydrogen bonds with the ligand are displayed. The hydrogen bonds themselves are shown as small green spheres in the main ADT window (figure 31).

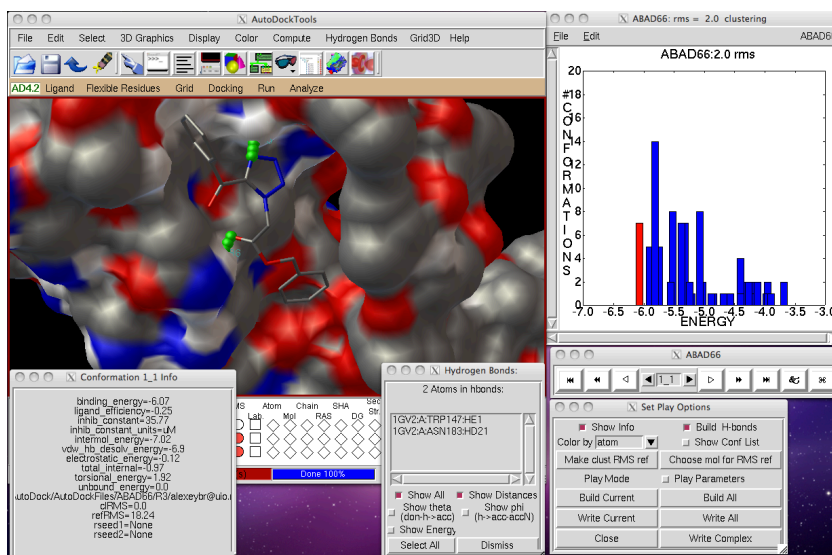


Fig. 31

There are many more tutorials and manuals that describe the process of docking, preparation of the GPF- and the DPF-files in more detail.²⁴ The preparation and the docking analysis of the docked compounds was done on a Macbook Pro mid 2010 model running OS X 10.6.8 with Python 2.5.

Chapter 4

Overview and pharmacological analysis of NSC87877 and NSC675460

In this part, a very general overview of the compounds NSC87877 and NSC675460 will be given. An in-depth look into the potential metabolic transformations of both compounds is provided.

4.1 NSC87877

Compound NSC87877 (figure 32) from the NCI (National Cancer Institute) library was identified by Martinez et al. as a potential inhibitor of the DBD in the c-Myb after virtual screening of the NCI library.⁶⁸

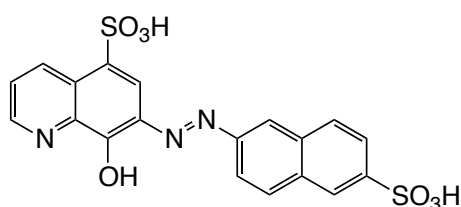


Fig. 32. NSC87877

The IUPAC name of the compound is (*E*)-8-hydroxy-7-((6-sulfonaphthalen-2-yl)diazenyl)quinoline-5-sulfonic acid. It is a known compound that was synthesized in the late 50's – early 60's.²⁵ Originally this compound and related ones were used as indicators for metal ions in complexometric titrations.²⁶

Compound NSC87877 can be viewed as consisting of two parts – the 8-hydroxyquinoline-5-sulfonic acid and the 6-aminonaphthalene-2-sulfonic acid – that are connected by an azo-bridge. Chemical compounds that have two large, fused aromatic systems (not always) on each end of an azo-group are often called azo-dyes. Most azo-dyes contain one azo-group, but there are numerous examples of dyes that have two or three azo-groups. The azo-group itself is not stable unless it is a part of an aromatic system that stabilizes it through electron delocalization. The azo-dyes can display a vast array of colors by changing the substituents on the diimide group. The cost of production of the azo-dyes are rather low, a factor that makes them attractive colorants in the industry.^{27, 41, 43}

During the past two decades, concerns have been expressed related to the use and production of azo-colorants. Dyes are used in the industry on an everyday basis to dye different textiles, papers, plastics, leathers and other products. As a matter of fact, like from any industrial plant, some dyes are released

into the environment. For some dyes, starting materials for the production of the colorants, especially ones based on benzidine analogues and biotransformation products were shown to be carcinogenic or mutagenic and were subsequently banned.²⁷

4.2 Metabolism of drugs

The material in this section, sections 4.3 and 4.4 is based on the reference.¹¹

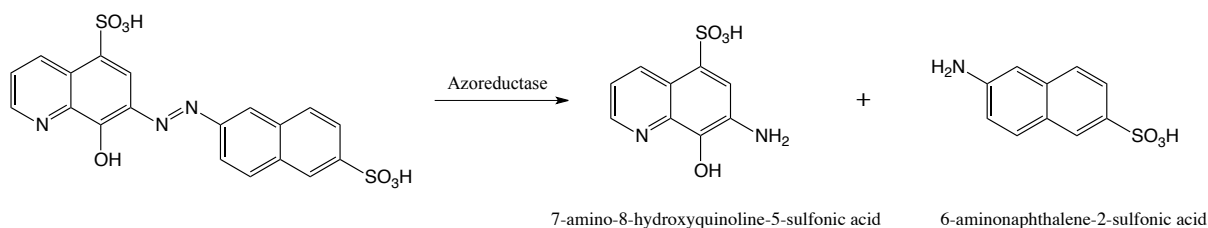
Pharmaceutical agents and other chemicals are foreign objects (xenobiotics) that the human body tries to deal with through a number of responses. Ideally, when a drug is given to a patient, it must reach its target site, cure the disease and leave the body without doing anything else. However, most drugs must undergo some sort of chemical reaction to aid excretion.

Drugs and other chemicals undergo a variety of transformations when they enter a human body. Enzymes that metabolize them are commonly located in liver, kidneys and lungs. Metabolism of drugs and chemicals also occurs in the gastrointestinal tract where the bacterial flora transforms xenobiotics into more / less harmful metabolites. Chemical reactions that occur during metabolism are usually divided into two big groups – phase I (P I) and phase II (P II) reactions.

The P I reactions are hydroxylation, reduction, oxidation and hydrolysis. Generally speaking, during the P I metabolism, new functional groups may be introduced / changed or exposed to P II pathways. The main idea behind P I processes is to make a molecule more polar so it can be excreted faster. The P II reactions can generally be called conjugation reactions – where an existing functional group (alcohol, phenol, amine etc.) is masked by the addition of a new group (acetyl, sulfate, amino acids etc.), thereby making the molecule even more polar. There are also chemicals/drugs that do not undergo metabolic reactions when they enter/go through the body – these compounds are usually excreted to a certain extent unchanged.

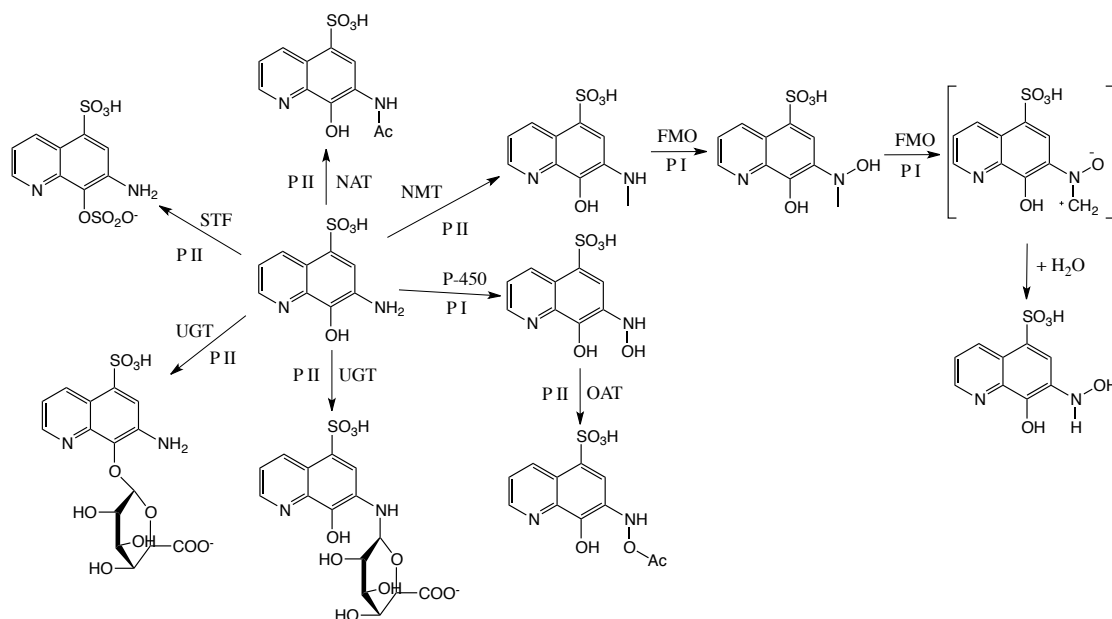
4.3 Pharmacological analysis of NSC87877

An analysis of potential metabolic pathways of compound NSC87877 begins with the identification of structural features in the molecule that might be substrates for P I/II reactions. The first and the most obvious feature, is the presence of the azo-bridge that is easily cleaved by the enzyme azoreductase, a NADPH-dependent system in the liver microsomes, converting the azo-bridge substituents into primary amines (P I reaction) (scheme 1):

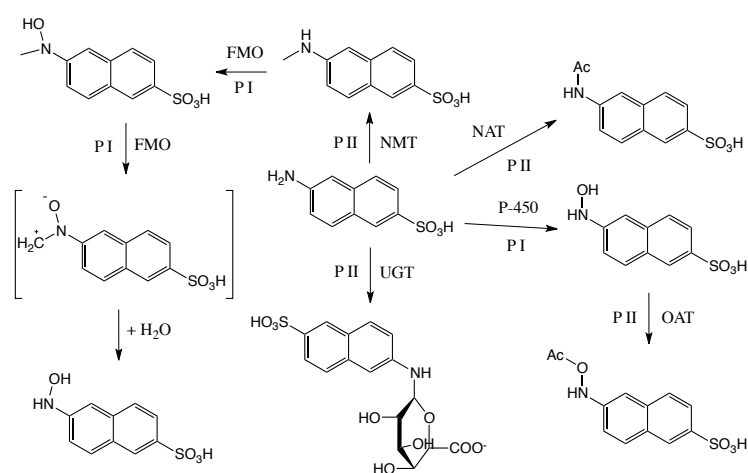


Scheme 1. Cleavage of the azo-bridge by the enzyme azoreductase

The products themselves can undergo further biotransformations (schemes 2 and 3):



Scheme 2. Theoretical metabolic pathways of the 7-amino-8-hydroxyquinoline-5-sulfonic acid.



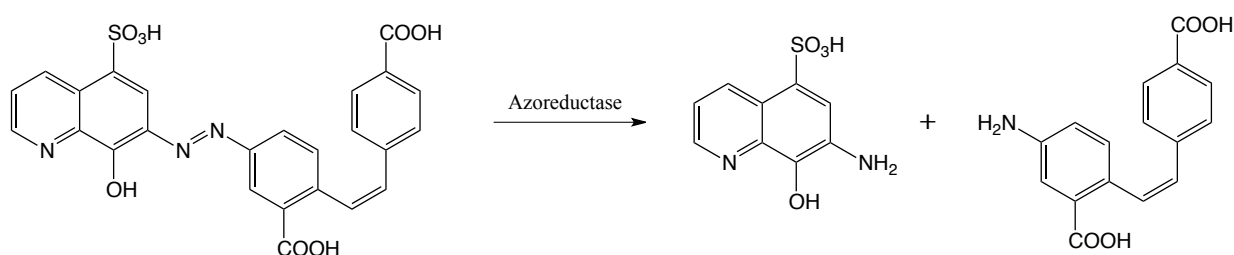
OAT: O – acetyltransferase; NAT: N – acetyltransferase; FMO: flavin – monoxygenase system; NMT: N – methyltransferase; P – 450: cytochrome CYP – 450 system; UGT: UDP – glucuronosyl transferase; STF: Sulfotransferase; P I: phase one reaction; P II: phase two reaction.

Scheme 3. Theoretical metabolic pathways of the 6-aminonaphthalene-2-sulfonic acid.

It is difficult to predict whether the different metabolic products will cause harm or detoxification. Despite the fact that there exists much knowledge about the metabolism of the drugs, there is no guarantee that these particular transformations will take place. On the basis of the notions from scheme 1, 2 and 3, one can assume that the NSC87877 will most probably lose its activity relatively fast once it enters the body. On the other hand, the way of administration of a drug can be subject to change and thus one can avoid certain metabolic pathways – an orally administered drug may undergo metabolic reactions in the gastrointestinal tract, while drugs administered intravenously avoid a given pathway. Another concern is the possibility of interactions between the pharmaceutical agent and metabolic products that can cause inhibition of a drug or serious side effects. Physiological and pathological factors such as age, heart condition etc. can as well result in different metabolic fates than the ones predicted – patients with a poorly functioning liver or kidneys may accumulate larger doses of toxic metabolites than healthy ones. Genetic differences may also give different rate of metabolism in certain pathways since some individuals may have different isoforms or even lack certain enzymes

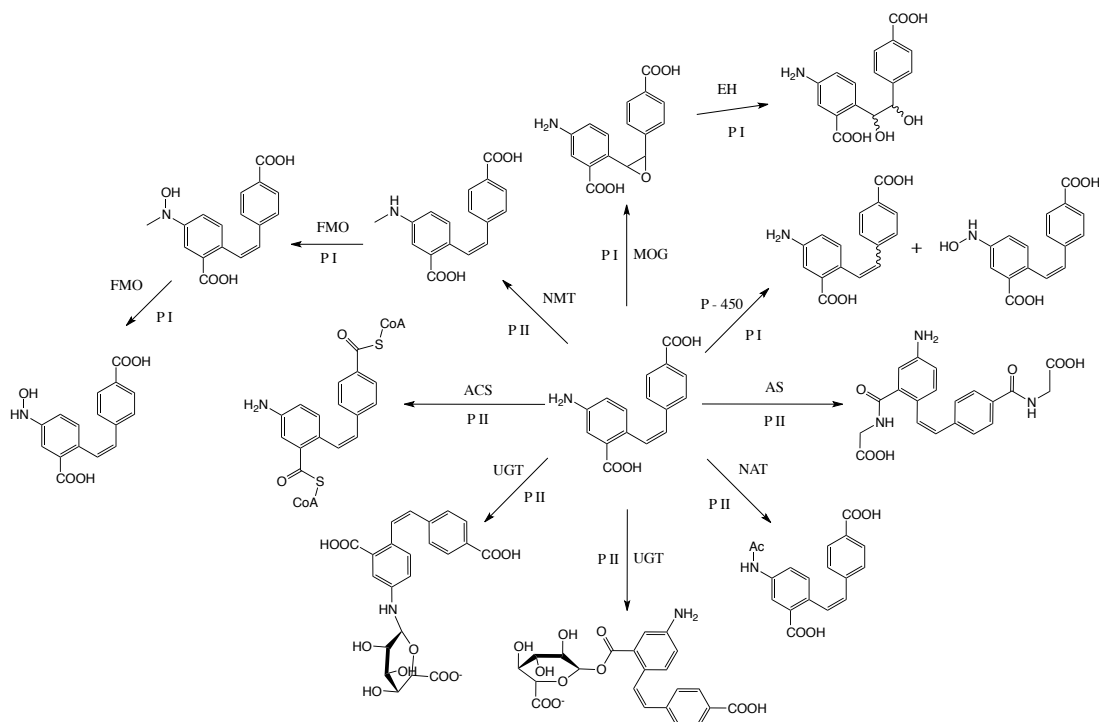
4.4 Pharmacological analysis of NSC675460

Martinez et al. discovered compound NSC675460 after screening of the NCI library for compounds that exhibited electronic similarities with NSC87877.⁶⁸ It had a more favorable energy of binding than other compounds similar to NSC87877. After an extensive search, it became clear that the compound exists only “virtually” in NCI databank. NSC675460 has several functional groups that most probably will undergo metabolic transformations inside human body (scheme 2,4 and 5):



Scheme 4. Cleavage of the azo-bridge by the enzyme azoreductase

All the proposed metabolic pathways for NSC675460 are purely theoretical and have not been corroborated. But one fact is clear – NSC675460 has many structural features that can serve as substrates for different enzymes present in the body. Most probably NSC675460 will lose its activity as a drug relatively fast due to conjugation pathways etc. One way to cope with this problem is to replace functional groups with biologically similar ones that do not undergo the same metabolic transformations. Investigation and docking analysis of different bioisosteres of NSC675460 was performed (for details see chapter 6, section 6.3 and 6.4).



NAT: N – acetyltransferase; FMO: flavin – monooxygenase system; NMT: N – methyltransferase; P – 450: cytochrome CYP – 450 system; UGT: UDP – glucuronosyl transferase; AS: Acyl synthetase; ACS: Acyl CoA synthetase; MOG: monooxygenase; EH: epoxide hydrolase; P I: phase one reaction; P II: phase two reaction.

Scheme 5. Theoretical metabolic pathways of the (Z)-5-amino-2-(4-carboxystyryl)benzoic acid

4.5 Lipinski's rule of 5

Lipinski's rule of 5²⁸ is four criteria that increase the chances of a drug to be highly soluble, cell permeable and thus, orally-active:

- Molecular weight < 500
- logP < 5
- Number of hydrogen donors < 5
- Number of hydrogen acceptors < 10

The criteria are based on calculated properties. The criteria have been established after experimental observations, for instance, studies have shown that a larger molecular weight of a compound results in poorer intestinal and blood brain barrier permeability.²⁹ The lipophilicity of a compound is expressed as a ratio (log P) of octanol solubility to aqueous solubility. The higher the log P value is, the more lipophilic and thus, less hydrophilic the compound is. An excessive number of hydrogen bonds in a compound harm the permeability across a membrane bi-layer. There are orally active therapeutic categories that violate the rule of 5 such as antibiotics, antifungals, vitamins and cardiac glycosides.

Chapter 5

The docking studies of NSC87877 and its analogues

In this chapter, the docking analysis of the five analogues of NSC87877 that were about to be synthesized will be presented. The synthesis of the five analogues will be presented in chapter 9.

5.1 Docking studies of NSC87877 and its analogues against the R2 region of the DBD in the c-Myb

It was decided to synthesize a series of the analogues of NSC87877 with an azo-bridge and a sulfonic acid group placed at different positions in a naphthalene system. The decision of which analogues were synthesized was based purely on the availability of the starting material – X-aminonaphthalen-Y-sulfonic acid – with an amino and a sulfonic acid groups in the different positions (figure 33). A docking analysis of NSC87877 and the analogues was performed. Special attention was drawn to how significant the sulfonic acid groups, the hydroxyl groups etc. are for binding with the DBD in c-Myb.

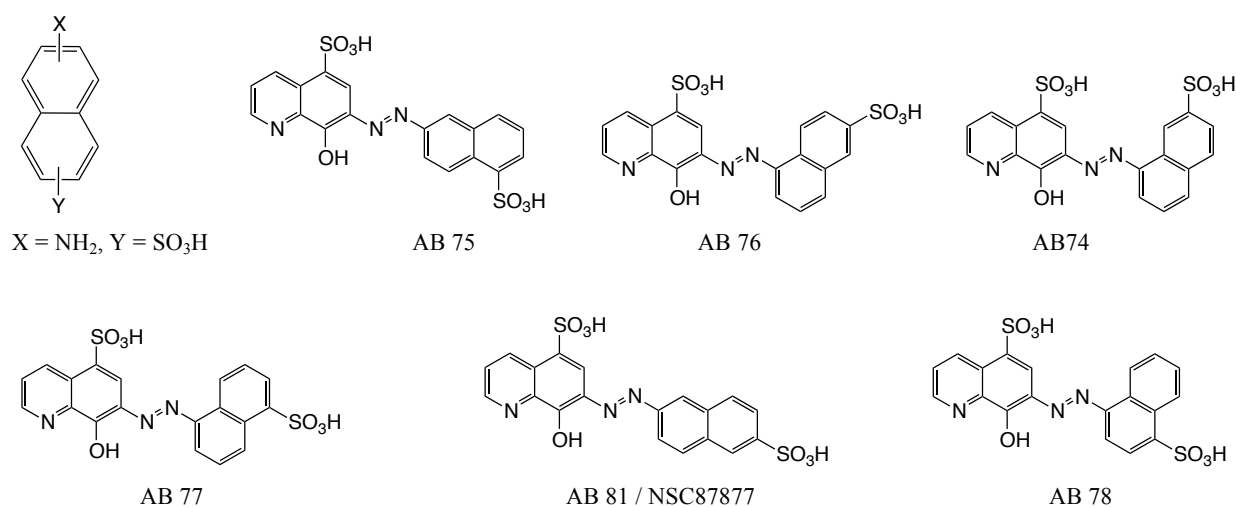


Fig. 33. Structures of the six compounds that were docked against the DBD in c-Myb.

From previous studies it was known that the sulfonic acid group in NSC87877 makes hydrogen bonds with the residues involved in the DNA binding.⁶⁸ Assuming that it was true, an *in silico* experiment in which the ligand molecule lacks functional groups capable of making hydrogen bonds, will most probably result in a less favourable free energy of binding and a higher inhibition constant. Both NSC87877 and the rest of the analogues have two sulfonic acid groups, an sp²-hybridized nitrogen in a quinoline-like moiety, a hydroxyl group and an azo-bridge that can function both as hydrogen donors

and acceptors. Removal of one or several functional groups must affect the binding energy etc. The binding pose and the location of the ligand will probably differ after deletion of a hydroxyl group etc.

Compounds NSC87877 and AB74 – 78 were docked against the residues Lys128 and Glu132 that are involved in hydrogen bonding with the c-Myb recognition element in the DNA. All chemical structures were drawn using JChemPaint, converted to the PDB-format using Avogadro and geometry optimized using iBabel, ff MMFF 94s force field, 250000 iterations with the steepest decent. The size of the grid box was 50 x 50 x 60, the grid spacing was default (0.375Å) and the grid box was centered on the Lys128 (22.153, 13.581 and 14.192). The search parameters were the following: 100 GA runs, 150 in population size, 5 million energy evaluations, translation step 1.0 Å, elitism 1, mutation and crossover rates of 0.02 and 0.8 respectively, and a local search rate of 0.06. The final docked conformations were clustered using a root-mean-square deviation of 2.0 Å. The results are summarized in table 1, while the complete table can be found in attachment 3, table 11.

Name of the compound	Free en. of bind. (kcal/mol)	R2 - Ki	Name of the compound	Free en. of bind. (kcal/mol)	R2 - Ki
AB81	- 9.91	53.99 nM	AB76	- 9.32	148.38 nM
AB81-1	- 8.45	645.09 nM	AB76-1	- 8.38	715.6 nM
AB81-2	- 9.25	165.66 nM	AB76-2	- 9.78	68.02 nM
AB81-3	- 8.47	620.97 nM	AB76-3	- 8.31	805.19 nM
AB81-4	- 9.99	47.68 nM	AB76-4	- 10.03	44.45 nM
AB81-5	- 9.68	79.66 nM	AB76-5	- 10.03	44.38 nM
AB74	- 10.94	9.5 nM	AB77	- 9.9	55.67 nM
AB74-1	- 8.91	296.19 nM	AB77-1	- 7.74	2.14 μM
AB74-2	- 10.27	29.84 nM	AB77-2	- 8.68	437.63 nM
AB74-3	- 8.47	622.47 nM	AB77-3	- 8.12	1.11 μM
AB74-4	- 10.6	17 nM	AB77-4	- 8.96	268.58 nM
AB74-5	- 10.85	11.12 nM	AB77-5	- 9.95	50.75 nM
AB75	- 9.08	221.08 nM	AB78	- 9.97	48.92 nM
AB75-1	- 8.97	266.27 nM	AB78-1	- 8.15	1.07 μM
AB75-2	- 8.68	436.7 nM	AB78-2	- 9.12	204.84 nM
AB75-3	- 8.6	497.86 nM	AB78-3	- 8.29	835.69 nM
AB75-4	- 9.31	151.1 nM	AB78-4	- 10.05	43.27 nM
AB75-5	- 9.18	187.14 nM	AB78-5	-10.04	43.45 nM

Table 1. Docking of NSC87877 and AB74 – 78 against the R2 region of the DBD in c-Myb. Ki stands for inhibition constant.

Compounds AB74, 75, 76, 77, 78 and AB81 showed similar binding energies upon docking to the R2 region of the DBD in c-Myb. Keeping in mind that the standard error in docking is around 2 – 3 kcal/mol, the results can't point which of the analogues would work better in a biological assays.

According to the literature, compounds with energies lower than -7 kcal/mol will most probably be specific, strong binding ligands – all of the analogues fall into that category.³⁰ Analysis of the top 5 conformations of every analogue showed that they bind to the two specific areas in the R2 region (figure 34).

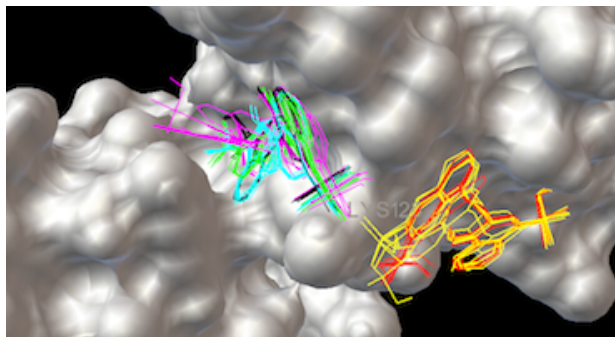


Fig. 34. The target molecule R2R3 is visualized with molecular surface (vdW). Residue Lys128 is marked. Top 5 conformations of AB 81 is green, AB74 – red, AB75 – purple, AB76 – yellow, AB77 – turquoise, AB78 – black. Image generated with ADT 1.5.4.

Electronic factors in the target molecule made one site particularly favorable for AB74 and 76. If this result would mean anything for the experimental analysis of *in vitro* activity was not clear. The top 5 conformations of AB75, 77, 78 and 81 were all docked inside some sort of a pocket. From one point of view it is a goal of the docking process to be sure that the ligand binds to the target in a pocket.

AutoDock, as it was mentioned before, treats the target molecule as a rigid structure although it is a common assumption that this is false. Proteins and macromolecules are dynamic structures that undergo conformational changes due to molecules that interact with them, solvent effects or interchanging hydrogen bonds. Fluctuations in the electronic environment make it impossible to predict with high degree of accuracy where the ligand will bind or not. The solvent molecules may impede the process of binding by more strongly solvating functional groups that could make hydrogen bonds with a ligand. Based on these notions, the binding site of AB75, 77, 78 and 81 is favored from one point of view, yet from another, some uncertainty regarding the dynamical behavior of the substrate must be addressed.

An investigation of how the elimination of the different functional groups in AB74 - 78 and 81 will affect the free energy of binding showed some general trends. Upon elimination of the sulfonic acid group from either the naphthalene or the quinoline moieties, the free energy of binding increased across all the analogues (approximately from -9 kcal/mol to -8 kcal/mol). Factors such as the absence of hydrogen bonds, electrostatic interactions between the sulfonic acid group and the residues in the R2 region may account for the trend. Elimination of the hydroxyl group showed a similar tendency for the same reasons. Substitution of the azo-bridge with a *trans* carbon-carbon double bond did not give any significant change in the free energy, neither substitution of the sp^2 -nitrogen in the quinoline moiety.

The loss of the free electron pair on the nitrogen atom may account for a very small drop in the free energy in some analogues.

All analogues including those missing hydroxyl group etc. showed the same trends in their binding pattern. Majority of the analogues bind exactly in the same region as AB74 – 78 and 81. The top scoring conformation of the compounds AB81, 81-1 – 81-5 bind in the same place (figures 35, 36). Upon visualization of the docked conformations of AB81, 81-1 – 81-5, it becomes clear that the substrates bind to the DBD in the two specific regions (figure 37).

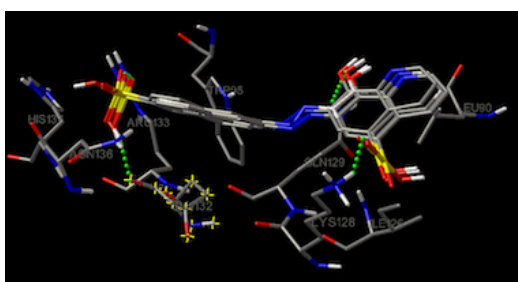


Fig. 35. AB81 – 81-5 bind at the same place. Hydrogen bonds are visualized with small green dots.

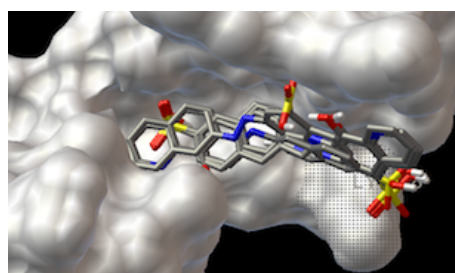


Fig. 36. Visualization of AB81 – 81-5 on the VdW surface of the R2 and R3.

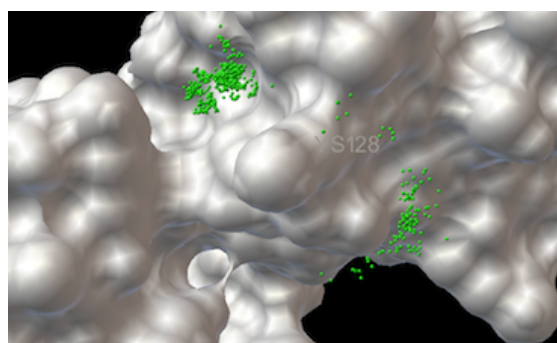


Fig. 37. 600 docked conformation of AB81 – 81-5 in the R2 region represented with small green spheres.

This indicates that the chemical skeleton (position of the carbon or nitrogen bridge may vary) of the analogues - (*E*)-7-(2-(naphthalen-2-yl)vinyl)quinoline – has a geometric and electronic structure that make it possible to induce hydrophobic interactions in these regions. Similar conclusions can be drawn from the investigation of AB74 – 78 and its analogues missing functional groups (attachment 4, p. 175).

5.2 Docking studies of NSC87877 and its analogues against the R3 region of the DBD in c-Myb

The compounds NSC87877 and AB74 – 78 were docked against the residues Lys183, Ser187, Asn183, Asn179 and Asn186 that are involved in hydrogen bonding with the c-Myb recognition element in DNA. All the chemical structures were drawn using JChemPaint, converted to the PDB-format using

Avogadro and geometry optimized using iBabel, ff MMFF 94s force field, 250000 iterations with the steepest decent. The size of the grid box was 74 x 44 x 40, the grid spacing was default (0.375Å) and the grid box was centred on Asn183 (-0.503, 15.076 and 7.389). The search parameters were following: 100 GA runs, 150 in population size, 5 million energy evaluations, translation step 1.0 Å, elitism 1, mutation and crossover rates of 0.02 and 0.8 respectively, and local search rate of 0.06. The final docked conformations were clustered using a root-mean-square deviation of 2.0 Å. The results are summarized in table 2 and the complete tables can be found in attachment 3, table 12.

Name of the compound	Free en. of bind. (kcal/mol)	R3 - Ki	Name of the compound	Free en. of bind. (kcal/mol)	R3 - Ki
AB81	- 6.63	13.73 µM	AB76	- 7.5	3.19 µM
AB81-1	- 7.75	2.08 µM	AB76-1	- 7.71	2.23 µM
AB81-2	- 6.74	11.56 µM	AB76-2	- 7.12	6.03 µM
AB81-3	- 6.83	9.87 µM	AB76-3	- 7.11	6.12 µM
AB81-4	-6.9	8.69 µM	AB76-4	- 7.6	7.71 µM
AB81-5	- 6.61	14.27 µM	AB76-5	- 7.65	2.47 µM
AB74	-7.23	5 µM	AB77	- 7.85	1.77 µM
AB74-1	- 7.06	6.73 µM	AB77-1	- 7.19	5.14 µM
AB74-2	- 7.09	6.36 µM	AB77-2	- 6.88	9 µM
AB74-3	- 6.76	11.09 µM	AB77-3	- 6.87	9.2 µM
AB74-4	- 7.11	6.09 µM	AB77-4	- 6.87	9.17 µM
AB74-5	- 7.04	6.86 µM	AB77-5	- 7.84	1.8 µM
AB75	- 6.49	17.35 µM	AB78	- 6.81	10.24 µM
AB75-1	- 7.9	1.61 µM	AB78-1	- 7.64	2.52 µM
AB75-2	- 6.29	24.68 µM	AB78-2	- 7.71	2.21 µM
AB75-3	- 6.47	18.21 µM	AB78-3	- 7.21	7.35 µM
AB75-4	- 6.63	13.77 µM	AB78-4	- 7.03	6.99 µM
AB75-5	- 6.79	10.47 µM	AB78-5	- 6.93	8.36 µM

Table 2. Docking of NSC87877 and AB74 – 78 against the R2 region of the DBD in c-Myb. Ki stands for inhibition constant.

The same trends continue to appear in the R3 region also. The free energy of binding is less favorable than in the R2 region, equally so with the binding affinity values. Elimination of the functional groups causes identical effects as in the R2 region. The top conformations of AB81 – AB81-5 all bind in the same place in the R3 and make hydrogen bonds with the residues involved in interactions with DNA (figure 38). Visualization of all conformations of AB81 – 81-5 shows that all the molecules were docked into a small area (figure 39). The top 5 conformations of AB74, 75, 76, 77, 78 and 81 all showed

a similar pattern in the docking (figure 40). Better visualizations of the docked conformations of AB74 - 78 and AB81 in the R3 region can be found in attachment 4, table 12.

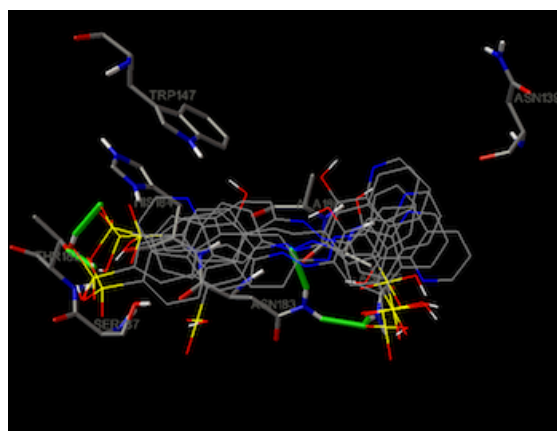


Fig. 38. Visualization of top scoring conform. of AB81 – AB81-5.

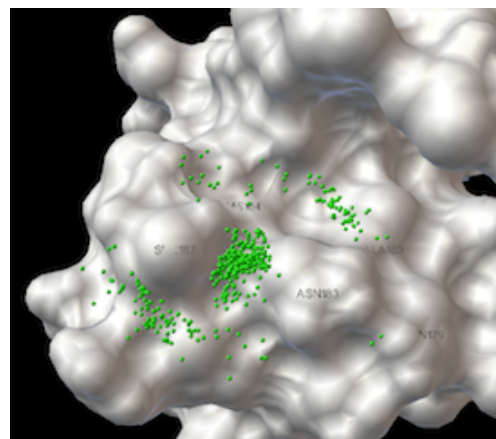


Fig. 39. 600 docked conformation of AB81 – AB81-5.

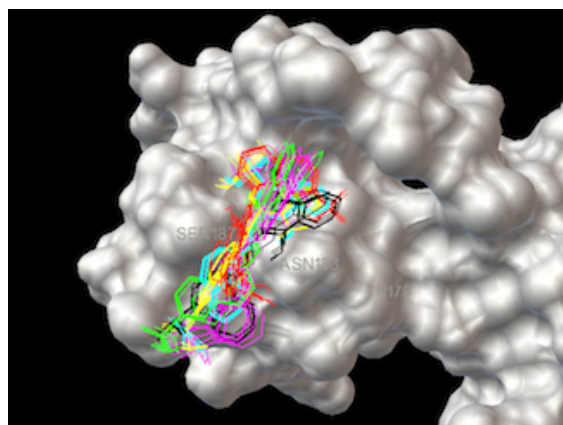


Fig. 40. The R3 region of the DBD in c-Myb. Top 5 conformations of AB81 is colored green, AB74 – red, AB75 – purple, AB76 – yellow, AB77 – turquoise, AB78 – black. Image was generated with ADT 1.5.4.

The elimination of the functional groups capable of making hydrogen bonds in compounds AB74 – 78, 81 resulted in less favorable free energy of binding between the compounds and the DBD in c-Myb. The trend shows the significance of the sulfonic acid and hydroxyl groups in induction of the favorable interactions between the target and the drug molecules.

From the docking studies of AB81 and its 5 analogues it became clear that these compounds might also be potent inhibitors of the DBD in c-Myb. From the previous work conducted by Martinez et al., it was known that compound NSC87877 had an IC₅₀ value of 1 μ M.⁶⁸ Theoretically, the compounds similar to NSC87877, that had better *in silico* results, could have even lower IC 50 values.

Computational studies of NSC675460 and its analogues

In this chapter, the bioisosteric approach in drug discovery will be discussed. The same methodology will be used in the modification of compound NSC675460. Docking studies of the designed compounds against the R2 and R3 region of the DBD in c-Myb will also be discussed.

6.1 Docking of NSC675460 against the R2 and R3 region in the DBD in c-Myb

Compound NSC675460, discovered by Martinez et al., was re-docked against the R2 and R3 region in the DBD of c-Myb to verify the previously obtained results.⁶⁸ The free energies of binding differed from the ones obtained by Martinez et al. (table 3).

Compound	Results from Martinez		Results from this study	
	R2	R3	R2	R3
NSC675460	-10.98 kcal/mol	- 8.52 kcal/mol	- 9.19 kcal/mol	- 6.57 kcal/mol

Table 3. Comparison of the docking results.

It is unclear how Martinez et al. obtained such good results, but this will not be discussed further. Visualization of the top 5 conformations of NSC675460 in the R2 and R3 regions gives the following picture (figures 41 – 44).

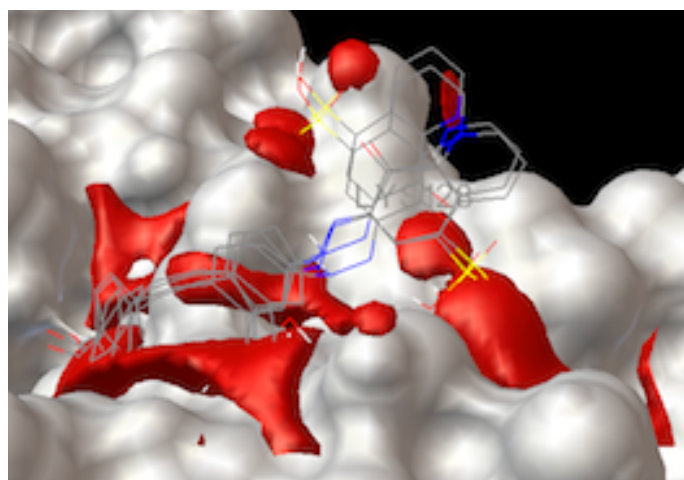


Fig 41. Residue Lys128 is marked with text. Red areas represent areas on the surface of the R2 region that have affinity for oxygen. The sulfonic and carboxylic acid groups are buried inside regions with oxygen affinity. It explains the big energy of binding.

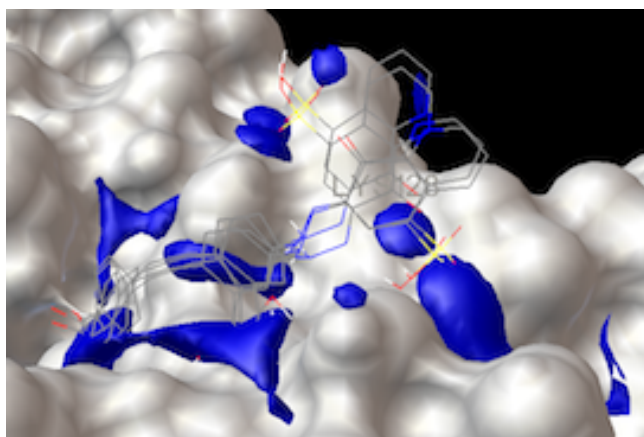


Fig 42. Residue Lys128 is marked with text. Blue areas represent areas on the surface of the R2 region in the DBD that have affinity for nitrogen. The same areas with affinity for oxygen, also have affinity for nitrogen

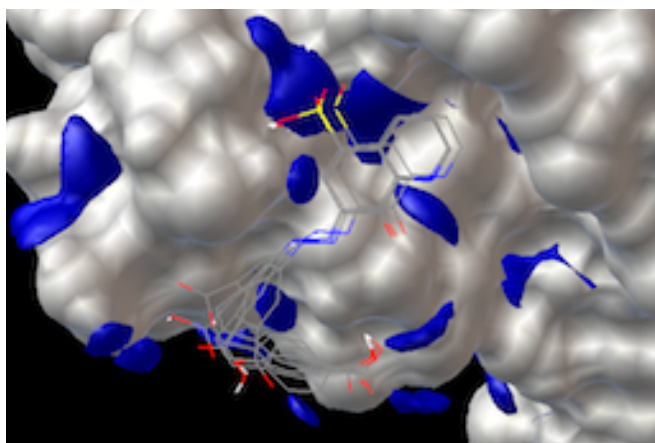


Fig. 43. The nitrogen affinity map in the R3 region of the DBD in the c-Myb.

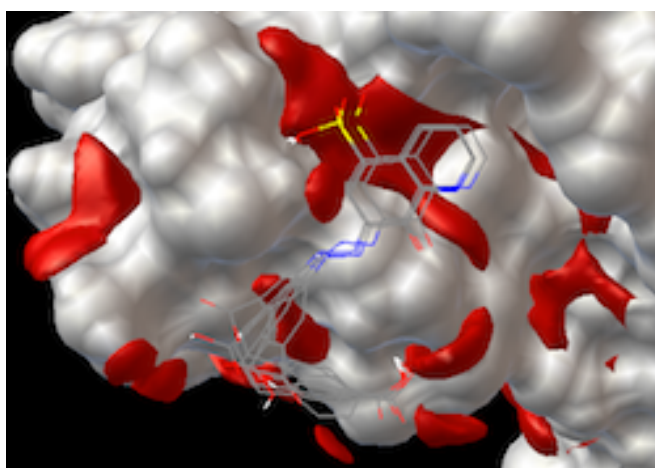


Fig. 44. Oxygen affinity map in the R3 region of the DBD in c-Myb. One can see that only the sulfonic acid group is buried inside the area with affinity for oxygen, the carboxylic acid groups do not make contact with these areas. It explains the quite small energy of binding of NSC 675469 in the R3 region.

The most important piece of information from figures 41 – 44 is that the same areas on the surface of the DBD in c-Myb have affinity both for oxygen and nitrogen. Theoretically, if the sulfonic or

carboxylic acid groups are substituted with other functional groups that bear nitrogen atoms instead of the oxygens, they will bind to the same area. This substitution may result in more favorable free energies of binding and give a more “potent” *in silico* drug candidate. This theory was investigated further.

6.2 The bioisosteric approach in the discovery of novel potential drugs

As was mentioned before, compound NSC675460 may undergo metabolic transformations. Its time of action as an active drug might be very short and will not lead to the desired effect. One way to deal with this type of concerns is to replace the functional groups in the molecule with similar ones that do not undergo metabolic degradation that easily. The substitution may also modify the toxicity, the chemical properties and the activity of the molecule of interest. A common term used about the substitution of functional groups with other groups presenting similar physical and chemical properties, is referred to as an isosteric replacement. The concept of chemical isosterism was first developed and described by Langmuir in 1919.³¹ It was used to describe physical similarities between atoms, molecules and functional groups. Langmuir’s theory was primarily based on the fact that atoms with the same number of valence electrons, from the same column in the periodic table had similar physicochemical properties. This idea was limited only to some atoms and molecules and was later developed further by Grimm and Hinsberg.¹¹

Substitution of functional groups and atoms in biologically active molecules is not simple since a drug molecule usually has a lot of different interactions with a target molecule/biological system that cannot be analyzed straightforwardly. The physicochemical properties of functional groups are somehow related to biological properties, but it is not easy to generalize this comparison across biological systems – a successful substitution in one system can give an adverse effect in another. The concept itself – bioisosterism – was pioneered by Friedman.¹¹ Later, Burger expanded the original definition by saying that: “Bioisosteres are compounds or groups that possess near equal molecular shapes and volumes, approximately the same distribution of electrons, and which exhibit similar physical properties such as hydrophobicity. Bioisosteric compounds affect the same biochemically associated systems as agonist or antagonists and thereby produce biological properties that are related to each other”.³²

Bioisosteric compounds are commonly divided into two subclasses: classical and non-classical. Classical bioisosteres are functional groups that obey the guidelines presented by Langmuir and Grimm. Non-classical bioisosteres do not follow the same rules as classical ones and often have different numbers of atoms than the group(s) they replace. Examples of classical bioisosteric replacement are substitution of a hydrogen atom with a fluorine or substitution of an amino group with a hydroxyl group.³³ An example of a non-classical replacement can be found in the development of the drug

isoproterenol (figure 45).¹¹ The phenolic hydroxyl and the sulfonamido group do not have same number of atoms, electrons or size. But they are both capable of donating a hydrogen atom and interacting with the target receptor as anions or as hydrogen bond donors. The groups also have roughly the same pKa. Another crucial factor is that the sulfonamide group is not metabolized by the O-methyltransferase, thereby increasing the time of action of the drug.

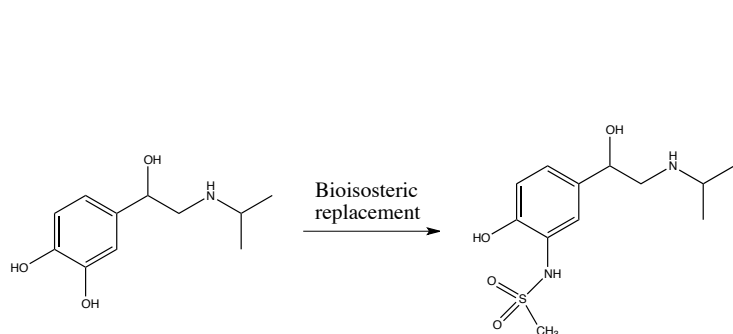


Fig. 45

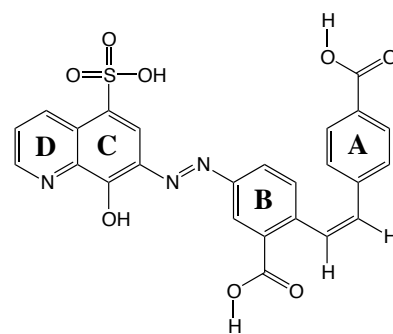


Fig. 46

6.3 Bioisosteric replacement of the functional groups in NSC675460

To simplify the discussion about the bioisosteric replacement of the existing functional groups in NSC675460, the aromatic rings will be named as it is depicted in the figure 46. One of the first and most obvious replacements is the substitution of the azo-bridge with a C=C *trans* double bond between the rings C and B, to prevent cleavage by the enzyme azoreductase. A comparison of the geometries of the substrates with N=N or C=C bonds showed that the overall orientation of the functional groups has only undergone minor changes (figure 47 – 50). The chemical structures were drawn using JChemPaint, converted to the PDB-format using Avogadro and geometry optimized using iBabel, ff MMFF 94s force field, 250000 iterations with the steepest decent, visualized with Avogadro and the bond lengths and angles were measured using Avogadro.

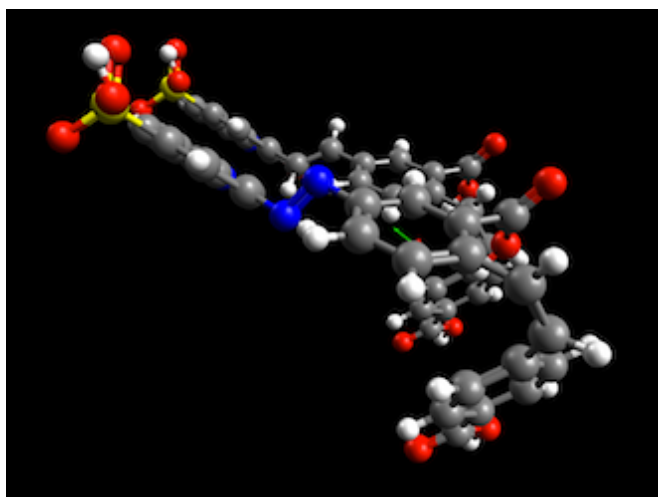


Fig. 47

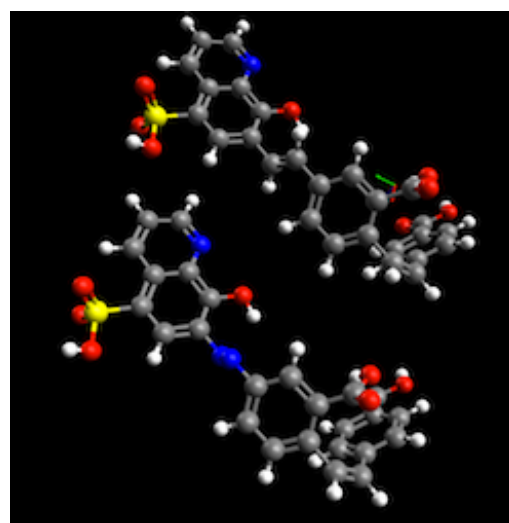


Fig. 48

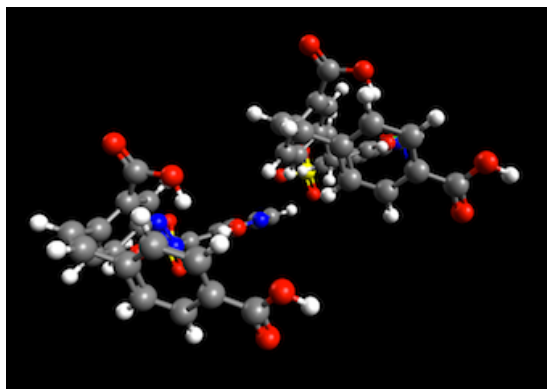


Fig. 49

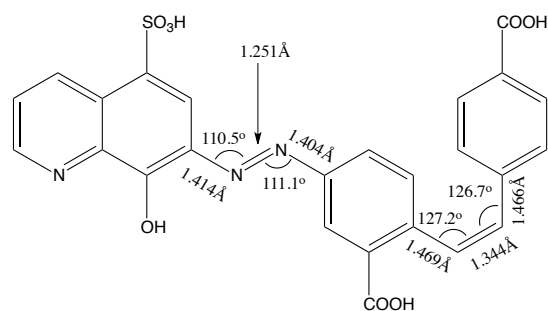
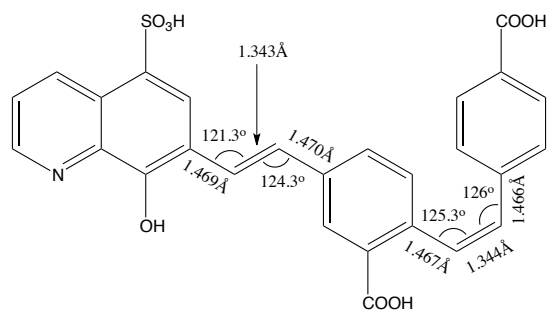


Fig. 50

6.4 Tetrazoles as bioisosteres of carboxylic acids

Substitution of the carboxylic acid group at the rings A and B with another functional group is the next step towards the design of a more robust molecule. 5-Substituted tetrazole rings are considered as bioisosteres of carboxylic acids – they often preserve or improve the biological activity of the parent drug.^{32, 34}

Tetrazole rings are 6 π -azopyrrole type ring systems that are usually classified as aromatic systems due to their reactivity. 5-Substituted tetrazoles are commonly referred to as tetrazolic acids if they have a free N-H bond. They exist as two tautomeric forms in nearly 1:1 ratio – *1H* and *2H* – that interconvert rapidly (figure 51).

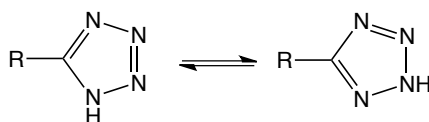


Fig. 51. Two tautomeric forms of the 5-substituted tetrazoles.

The acidity of the N-H is largely dependent on the substituent at the 5'-position. Both 5-alkyl- and 5-aryltetrazoles have a pK_a that is similar to that of carboxylic acids due to their ability to stabilize negative charge by electron delocalization.³⁵ Tetrazoles, like carboxylic acids, are ionized at physiological pH and have a planar structure. Tetrazoles in their ionized state are more lipophilic than the corresponding carboxylic acids – which can make it easier for a drug molecule to penetrate the cell

membrane. Delocalization of a negative charge over the larger surface area of tetrazole rings, compared to carboxylic acids, may both enhance and diminish binding affinity of a drug containing a tetrazole ring. It is difficult to predict how the biological activity will vary upon substitution with the tetrazole moiety. Recent studies show that in the cases where a tetrazole ring improved *in vitro* results, it was due its capability to make hydrogen bonds with the residues at the target molecule.³⁶

Both the carboxylic and tetrazole groups may function as CYP-450 substrates for oxidative metabolic processes, but the tetrazole analogues do not react in the phase II conjugation pathways. Unfortunately it has been shown that 5-aryltetrazoles undergo conjugation reactions and form β -N-glucuronides (figure 52).³⁷

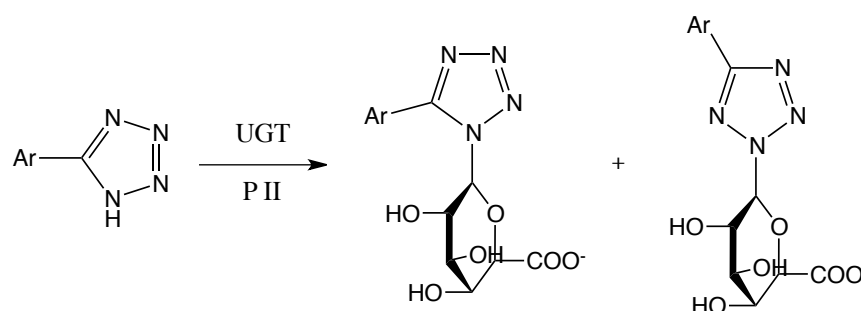
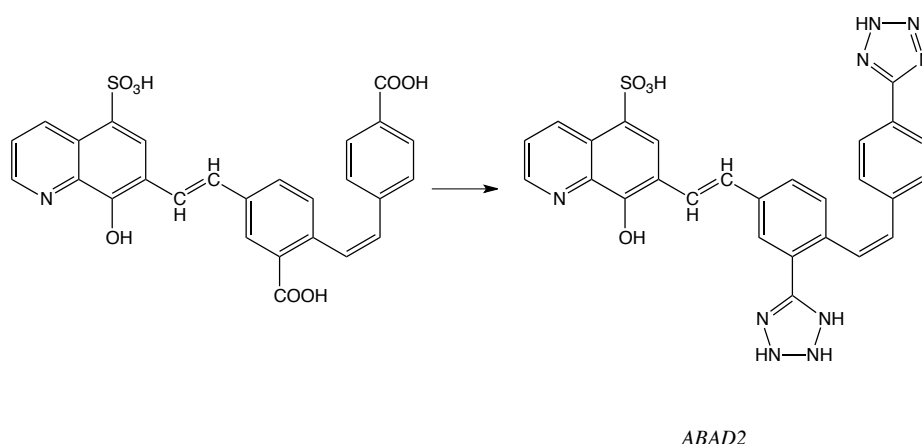


Fig. 52. The phase II conjugation reaction of tetrazoles.

The second step of the bioisosteric replacement process in NSC675460 consisted of the substitution of the carboxylic acid groups at rings A and B with tetrazole rings:



From the docking analysis of compound ABAD2 against the residues in the R2 and R3 region, promising results were obtained. In the R2 region: The free energy of binding: -11.79 kcal/mol; $K_i = 2.26$ nM. In the R3: The free energy of binding: -10.54 kcal/mol; $K_i = 18.64$ nM. Encouraged by these results, several (68) bioisosteres and analogues of compound ABAD2 were designed and docked against the residues in the R2 and R3 regions.

6.5 Docking of NSC675460 analogues against the R2 region of the DBD in c-Myb

Compounds ABAD1 – 68 were docked against residues Lys128 and Glu132 that are involved in the hydrogen bonding with the c-Myb recognition element in DNA (table 4). All chemical structures were drawn using JChemPaint, converted to the PDB-format using Avogadro and geometry optimized using iBabel, ffMMFF 94s force field, 250000 iterations with the steepest decent. The size of the grid box was 50 x 50 x 60, the grid spacing was default (0.375Å) and the grid box was centered on Lys128 (22.153, 13.581 and 14.192). The search parameters were the following: 100 GA runs, 150 in population size, 5 million energy evaluations, translation step 1.0 Å, elitism 1, mutation and crossover rates of 0.02 and 0.8 respectively, and local search rate of 0.06. The final docked conformations were clustered using a root-mean-square deviation of 2.0 Å. The results are summarized in table 4 and the complete tables can be found in attachment 3, table 13.

Name of the compound	Free en. of bind. (kcal/mol)	R2 - Ki	Name of the compound	Free en. of bind. (kcal/mol)	R2 - Ki
ABAD1	- 10.83	11.43 nM	ABAD35	- 7.66	2.43 μM
ABAD2	- 11.79	2.26 nM	ABAD36	- 7.91	1.6 μM
ABAD3	-8.17	1.03 μM	ABAD37	- 8.76	379.6 nM
ABAD4	- 10.22	32.19 nM	ABAD38	- 8.49	602.21 nM
ABAD5	- 7.98	1.41 μM	ABAD39	- 10.38	24.57 nM
ABAD6	- 8.07	1.22 μM	ABAD40	- 6.57	15.25 μM
ABAD7	- 5.58	81.62 μM	ABAD41	- 6.29	24.2 μM
ABAD8	- 8.1	1.15 μM	ABAD42	- 8.83	338.03 nM
ABAD9	- 8.01	1.35 μM	ABAD43	- 9.13	202.59 nM
ABAD10	- 11.2	6.17 nM	ABAD44	- 10.87	10.82 nM
ABAD11	- 9.07	226.45 nM	ABAD45	- 9.84	61.63 nM
ABAD12	- 10.78	12.52 nM	ABAD46	- 10.35	25.7 nM
ABAD13	- 7.41	3.67 μM	ABAD47	- 10.09	40.51 nM
ABAD14	- 8.67	438.4 nM	ABAD48	-10.01	45.79 nM
ABAD15	- 7.02	7.16 μM	ABAD49	- 6.68	12.59 μM
ABAD16	- 9.17	190.88 nM	ABAD50	- 8.5	301.42 nM
ABAD18	- 10.35	26.06 nM	ABAD51	-10.41	23.31 nM
ABAD19	- 3.8	1.63 mM	ABAD52	- 8.52	573.08 nM
ABAD20	- 8.25	903.44 nM	ABAD53	- 10.47	21.29 nM
ABAD21	- 10.84	11.32 nM	ABAD54	- 12.87	367.91 pM
ABAD22	- 8.29	285.43 nM	ABAD55	- 10.04	43.46 nM
ABAD23	- 9.61	89.95 nM	ABAD56	- 8.58	511.22 nM
ABAD24	- 9.55	100.22 nM	ABAD57	- 9.7	77.94 nM
ABAD25	- 8.83	335.53 nM	ABAD58	- 9.25	167.2 nM
ABAD26	- 7.17	5.58 μM	ABAD59	- 10.94	9.6 nM

ABAD27	- 9.51	107.16 nM	ABAD60	- 7.79	1.95 μ M
ABAD28	- 10.24	31.34 nM	ABAD61	- 7.4	3.77 μ M
ABAD29	- 8.92	290.06 nM	ABAD62	-4.84	282.95 μ M
ABAD30	- 8.61	491.93 nM	ABAD63	-10.27	29.69 nM
ABAD31	- 8.7	420.28 nM	ABAD64	-9.57	96.35 nM
ABAD32	- 9.87	58.17 nM	ABAD65	-10.31	27.87 nM
ABAD33	-8.89	306.97 nM	ABAD67	-8.12	1.12 μ M
ABAD34	- 9.3	151.41 nM	ABAD68	-8.16	1.05 μ M

Table 4. Docking of compounds ABAD1 – 68 against the R2 region in c-Myb. Ki stands for inhibition constant.

6.6 Docking of NSC675460 analogues against the R3 region of the DBD in c-Myb

Compounds ABAD1 – 68 were docked against residues Lys183, Ser187, Asn183, Asn179 and Asn186 that are involved in the hydrogen bonding with the c-Myb recognition element in DNA (table 5). All chemical structures were drawn using JChemPaint, converted to the PDB-format using Avogadro and geometry optimized using iBabel, ff MMFF 94s force field, 250000 iterations with steepest decent. The size of the grid box was 74 x 44 x 40, the grid spacing was default (0.375Å) and the grid box was centred on Asn183 (-0.503, 15.076 and 7.389). The search parameters were following: 100 GA runs, 150 in population size, 5 million energy evaluations, translation step 1.0 Å, elitism 1, mutation and crossover rates of 0.02 and 0.8 respectively, and local search rate of 0.06. The final docked conformations were clustered using a root-mean-square deviation of 2.0 Å. The results are summarized in table 5 and the complete tables can be found in attachment 3, table 14.

Name of the compound	Free en. of bind. (kcal/mol)	R3 - Ki	Name of the compound	Free en. of bind. (kcal/mol)	R3 - Ki
ABAD 1	- 10.16	35.98 nM	ABAD35	- 6.36	21.87 μ M
ABAD2	- 10.54	18.64 nM	ABAD36	- 6.01	39.27 μ M
ABAD3	- 7.76	2.05 μ M	ABAD37	- 8.54	552.15 nM
ABAD4	- 7.53	3 μ M	ABAD38	- 7.13	5.93 μ M
ABAD5	- 8.21	952.32 nM	ABAD39	- 8.69	426.76 nM
ABAD6	- 6.14	31.65 μ M	ABAD40	- 7.13	5.97 μ M
ABAD7	- 5.21	150.7 μ M	ABAD41	- 5.78	57.95 μ M
ABAD8	- 6.54	15.99 μ M	ABAD42	- 7.43	3.57 μ M
ABAD9	- 6.79	10.46 μ M	ABAD43	- 9.89	56.41 nM
ABAD10	- 10.26	30.21 nM	ABAD44	- 10.67	14.98 nM
ABAD11	- 5.36	118 μ M	ABAD45	- 11.47	3.9 nM
ABAD12	-10.56	18.06 nM	ABAD46	- 6.43	19.41 μ M
ABAD13	- 6.38	21.21 μ M	ABAD47	- 7.49	3.26 μ M
ABAD14	- 8.33	786.81 nM	ABAD48	- 6.83	9.8 μ M

ABAD15	- 6.35	22.34 μ M	ABAD49	- 6.9	8.8 μ M
ABAD16	- 6.97	7.76 μ M	ABAD50	- 9.12	207.19 nM
ABAD18	- 7.41	3.68 μ M	ABAD51	- 10.41	23.22 nM
ABAD19	- 4.1	992.9 μ M	ABAD52	- 9.97	48.82 nM
ABAD20	- 7.83	1.82 μ M	ABAD53	- 10.44	22.16 nM
ABAD21	- 9.15	197.17 nM	ABAD54	- 9.56	98.61 nM
ABAD22	- 5.7	58.3 μ M	ABAD55	- 9.89	55.88 nM
ABAD23	- 10.03	44.59 nM	ABAD56	- 8.56	532.06 nM
ABAD24	- 7.82	1.86 μ M	ABAD57	- 8.65	457.09 nM
ABAD25	- 6.61	14.21 μ M	ABAD58	- 8.2	972.63 nM
ABAD26	- 7.07	6.55 μ M	ABAD59	- 8.43	658.44 nM
ABAD27	- 8.93	282.6 nM	ABAD60	- 8.34	772.42 nM
ABAD28	- 10.52	19.42 nM	ABAD61	- 5.69	67.23 μ M
ABAD29	- 7.49	3.21 μ M	ABAD62	-4.64	398.44 μ M
ABAD30	- 9.18	186.68 nM	ABAD63	-7.14	5.85 μ M
ABAD31	- 7.53	3.02 μ M	ABAD64	-7.01	7.16 μ M
ABAD32	- 6.59	14.77 μ M	ABAD65	-7.78	1.97 μ M
ABAD33	- 6.42	19.55 μ M	ABAD67	-6.65	13.46 μ M
ABAD34	- 6.08	35.18 μ M	ABAD68	-7.82	1.87 μ M

Table 5. Docking of compounds ABAD1 – 68 against the R2 region in c-Myb. Ki stands for inhibition constant.

6.7 Re-docking

Compounds ABAD1, 2, 10, 12, 45, 51 and 53 showed the best binding affinities after docking against the residues in the R2 and R3 regions in the DBD in c-Myb. Every compound was re-docked in the same regions with the same parameters 25 more times, giving a total amount of docking runs per region per compound of 26. A short summary with average values of the free energies of binding and the inhibition constants is given in table 6 and the complete tables can be found in attachment 3, table 15.

Compound	Free energy of bind.	Ki R2 (nM)	Free energy of bind.	Ki R3 (nM)
	R2 (kcal/mol)		R3 (kcal/mol)	
ABAD1	Average: - 10.03	Average: 61.93	Average: - 10.05	Average: 56.02
	Std.: 0.51	Std.: 51.72	Std.: 0.48	Std.: 41.05
ABAD2	Average: - 10.79	Average: 17.66	Average: - 10.57	Average: 19.08
	Std.: 0.53	Std.: 15.03	Std.: 0.23	Std.: 7.17
ABAD10	Average: - 9.23	Average: 328.62	Average: - 8.98	Average: 503.27
	Std.: 0.77	Std.: 406.6	Std.: 0.65	Std.: 1012
ABAD12	Average: - 10.92	Average: 11.55	Average: - 10.33	Average: 30.77
	Std.: 0.31	Std.: 8.03	Std.: 0.31	Std.: 17.47
ABAD45	Average: - 10.17	Average: 42.32	Average: - 11.50	Average: 4.48
	Std.: 0.38	Std.: 26.96	Std.: 0.36	Std.: 2.8

ABAD51	Average: - 10.24	Average: 37.27	Average: - 10.45	Average: 23.40
	Std.: 0.37	Std.: 24.39	Std.: 0.31	Std.: 10.66
ABAD53	Average: - 10.59	Average: 27.14	Average: - 11.08	Average: 8.47
	Std.: 0.57	Std.: 28.52	Std.: 0.28	Std.: 4.49

Table 6. Re-docking of compound ABAD1, 2, 10, 12, 45 and 53 against the R2 and R3 regions in *c-Myb*.

Analysis of the binding poses of the compounds revealed some general trends. Taking the re-docking/statistical analysis of compound ABAD1 as an example, the 2600 conformations in the R2 and R3 regions can be visualized as in figures 53 and 54.

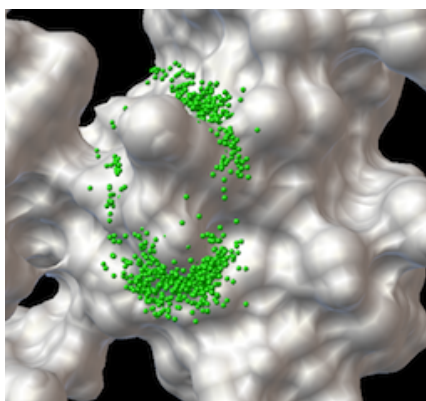


Fig. 53. 2600 conformations of ABAD1 docked in the R2.

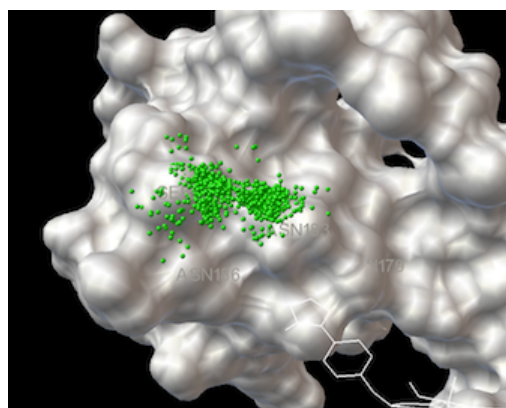


Fig. 54. 2600 conformations of ABAD1 docked in the R3.

As it can be seen from the figures 53 and 54, the conformations are grouped around residues Lys128, Ser187, Asn183 and Asn179 that are the primary target sites for the interactions/binding with the DNA chain. Another way to view the binding mode of ABAD1 is to visualize the top-ranking conformations from every single run (26 conformations at all) based on the binding affinities and the RMSD values (figures 55 and 56).

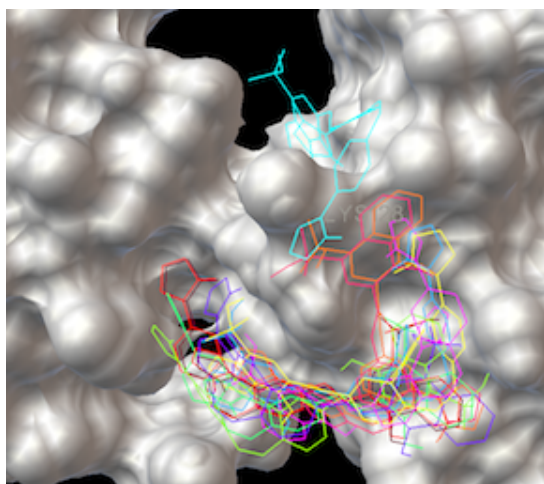


Fig. 55. The top-ranking conformations of the ABAD1 in the R2 region.

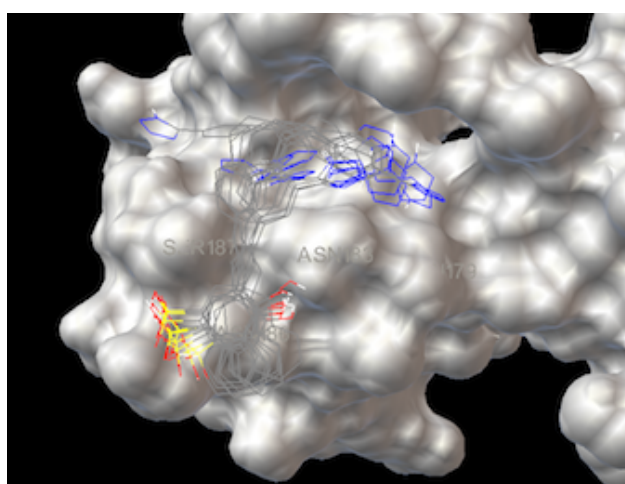


Fig. 56. The top-ranking conformations of the ABAD1 in the R3 region.

Based on these criteria, the compounds ABAD2, 45 and 53 were the best hits after analysis of the bioisosteres of compound NSC675460. An analysis of the top conformations of compound ABAD2 in the R2 region showed that the orientation of the conformations are similar – the sulfonic acid group and the tetrazole rings are placed in areas with high affinities for oxygen and nitrogen atoms respectively (figure 57). Analysis of all 2600 conformations of the compound ABAD 2 in the R2 region showed specific grouping around Lys128 residue (figure 58).

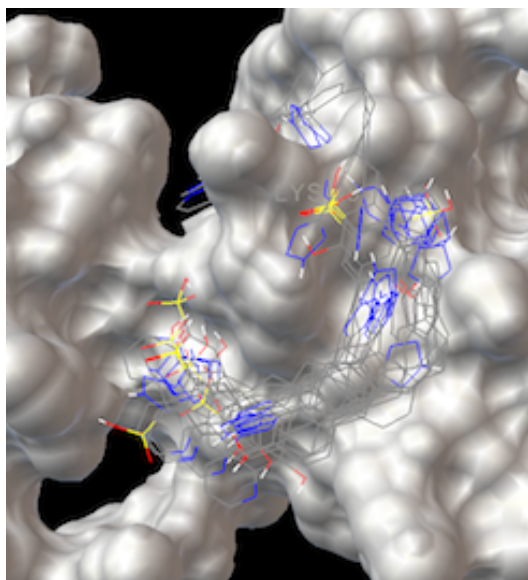


Fig. 57. Top conformation of the ABAD2 in the R2.

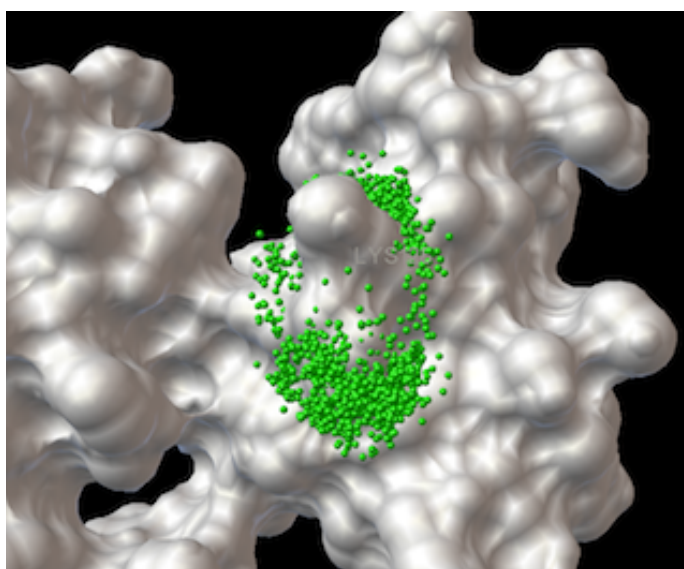


Fig. 58. 2600 conformation of ABAD2 in the R3 region.

The same conclusions can be drawn after analysis of the top conformations of ABAD2 in the R3 region (figures 59 – 61).

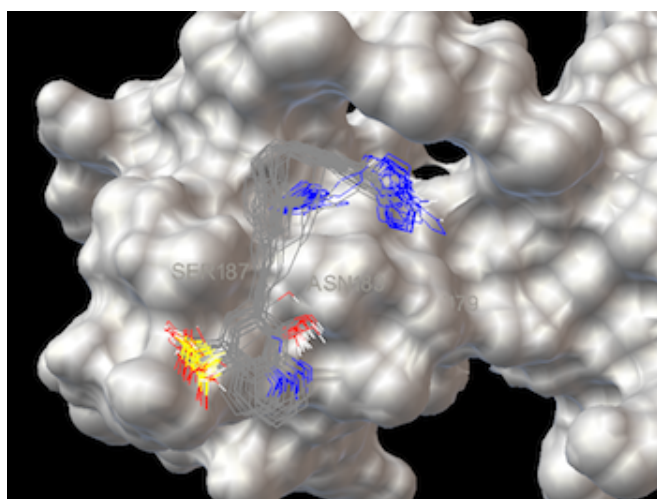


Fig. 59. Top conform. of ABAD2 from every run. Tetrazole rings and sulf. acid groups have the same orientation.

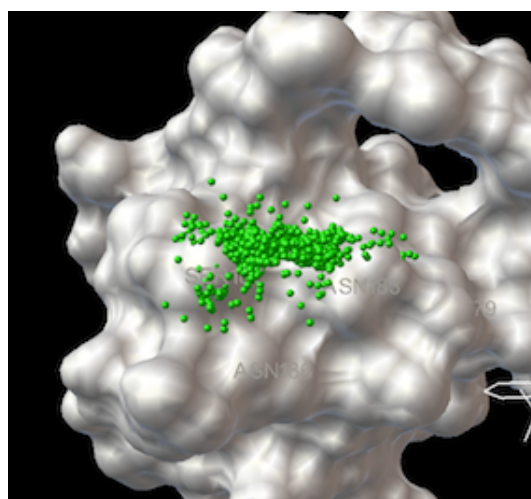


Fig. 60. All 2600 docked conformations of ABAD2 represented with small green spheres.

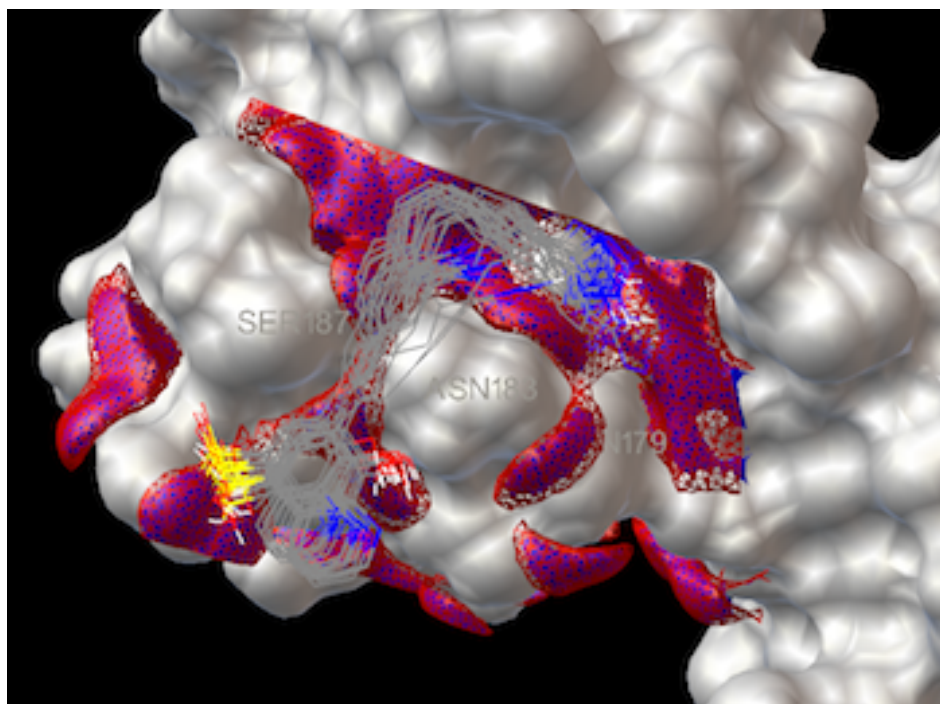


Fig. 61. Top conformations of ABAD2 from every single run (total 26 conformations). Red and blue areas represent regions on the surface of the DBD that have affinity for nitrogen and oxygen. Sulfonic acid groups and tetrazole rings are positioned right beside areas with high oxygen/nitrogen affinity.

Visualizations similar to these that show bindings/conformations of the compounds ABAD45 and ABAD53 can be found in attachment 4, p. 187. It is not possible to predict which of the three compounds will show the best results in biological testing since the affinity values lie very closely. One way to improve *in silico* results is to employ software programs such as CHARMM, GOLD, eHiTS etc. that use different scoring and docking functions than AutoDock 4, for the analysis of the top lead compounds.⁷⁴

Chapter 7

In silico analysis of the specificity of the compounds for the DBD in c-Myb

The docking of the selected compounds against the DBD of the Sp1 protein to probe for selectivity will be presented.

7.1 Docking of selected compounds against the ZF2/ZF3 in the DBD of the Sp1 protein

During the design of novel drug candidates, one of the main requirements is the specificity of a drug for a target. Interactions of a drug with other proteins or processes can lead to serious side effects and/or a decreased time of action as a drug. The aim of the project was to discover a compound that would bind only to the R2 and R3 in c-Myb and inhibit its DNA binding activity without interfering with other proteins. The DBD in the Sp1 protein was chosen as a domain for testing the specificity of compounds AB74 – 78, 81 and ABAD2, 45, 53. The DBD in the Sp1 protein consists of three regions called zinc finger 1, 2 and 3 (ZF 1, ZF 2 and ZF3). Only the ZF2/ZF3 play a major role in the binding to the DNA chain.³⁸ The selected compounds were docked against the ZF2/ZF3 (table 7).

Compound	Free en. of bind. ZF 2 (kcal/mol)	ZF 2 – Ki (nM)	Free en. of bind. ZF 3 (kcal/mol)	ZF 3 – Ki (nM)
AB 81	- 10.56	18.24 nM	- 9.85	59.99 nM
AB 74	- 9.79	66.99 nM	- 7.93	1.53 μ M
AB 75	- 8.95	274.5 nM	- 8.3	826.82 nM
AB 76	- 11.01	11.72 nM	- 9.6	92.57 nM
AB 77	- 10.29	28.84 nM	- 10.47	21.28 nM
AB 78	- 10.19	33.95 nM	- 8.8	310.86 nM
ABAD 2	- 12.16	1.22 nM	- 11.03	8.29 nM
ABAD 45	- 11.23	5.9 nM	- 10.87	10.83 nM
ABAD 53	- 10.71	14.14 nM	- 10.0	46.47 nM

Table 7. Docking of AB74 – 78, 81 and ABAD2, 45, 53 against the ZF2/ZF3 in the DBD of the Sp1 protein.

As it can be seen from the table, all values are below -7 kcal/mol, which is a typical threshold value for distinguishing molecules will likely bind to a target. Unfortunately this means that none of the compounds will bind *specifically* to the DBD in c-Myb – there will also be some degree of inhibition of the DBD in the Sp1 protein. On the other hand, the compounds were first docked against the ZF2 and then against the ZF3 – although it is known that the ZF1 – 3 exist as a single complex.³⁹ Factors as overlap of regions, a different 3D-structure when the three fingers exist in a complex etc. have not been

included, making the docking results not completely accurate. Compounds AB74 – 78, AB81 have been tested in the electrophoretic mobility shift assay against the DBD of the Sp1 protein and were found to inhibit it with IC 50 values in micromolar region – meaning that the docking results for the tested compounds predicted the experimental results relatively accurately.⁶⁹

Chapter 8

Biological testing of compounds AB74 – 78, AB81

In this chapter, the biological testing will be discussed briefly. The biological testing was done by members from the MYB group under the supervision of Professor Odd S. Gabrielsen at the Department of Molecular, University of Oslo. Comparison of the biological results and the *in silico* analysis will also be given. A detailed procedure for the tests that were performed has been previously published.⁶⁹

8.1 EMSA-analysis

EMSA – electrophoretic mobility shift assay – is a method for studying interactions between DNA and proteins. The method is very sensitive and makes it possible to detect and analyze very small amounts of protein. A solution of a protein that is suspected to bind to DNA, is mixed with a ³²P-labeled DNA sequence that contains a binding domain for that protein. Subsequently, the solution is subjected to electrophoresis through a polyacrylamide gel to separate a protein/DNA complex from the DNA strains that have not been bound by the protein. The molecules are separated according to their charge and size, so the protein/DNA complex will move slower than the unbound DNA because of its charge and size. The protein/DNA complex will be found closer to the cathode end; the free DNA will be located closer to the anode end (figure 62). The distribution of the different species is usually determined by an autoradiography technique.

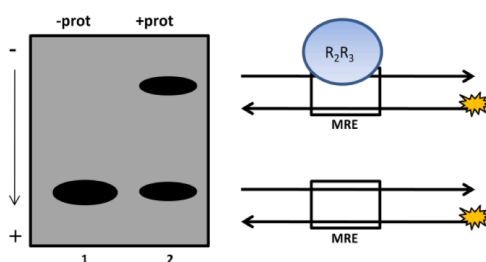


Fig. 62. Electrophoretic Mobility Shift Assay. The example shows the binding reaction between the R2/R3 regions of the DBD in *c-Myb* and ³²P labeled DNA sequence that binds to these two regions. The protein/DNA complex is separated from unbound DNA by electrophoresis (well 2). The upper band is the protein/DNA complex, while the lower one is the unbound DNA. To well 1 is added only DNA. The arrow shows the direction of migration of the molecules. Fig. is taken from ref. 69.

Every compound was first mixed with the R2 and R3 regions of the DBD in *c-Myb* and then added to the ³²P-labeled DNA sequence GCATTATAACGGTTT. Then, the EMSA has been performed in order to separate the complexes from the unbound molecules. The results can be found in figure 63.

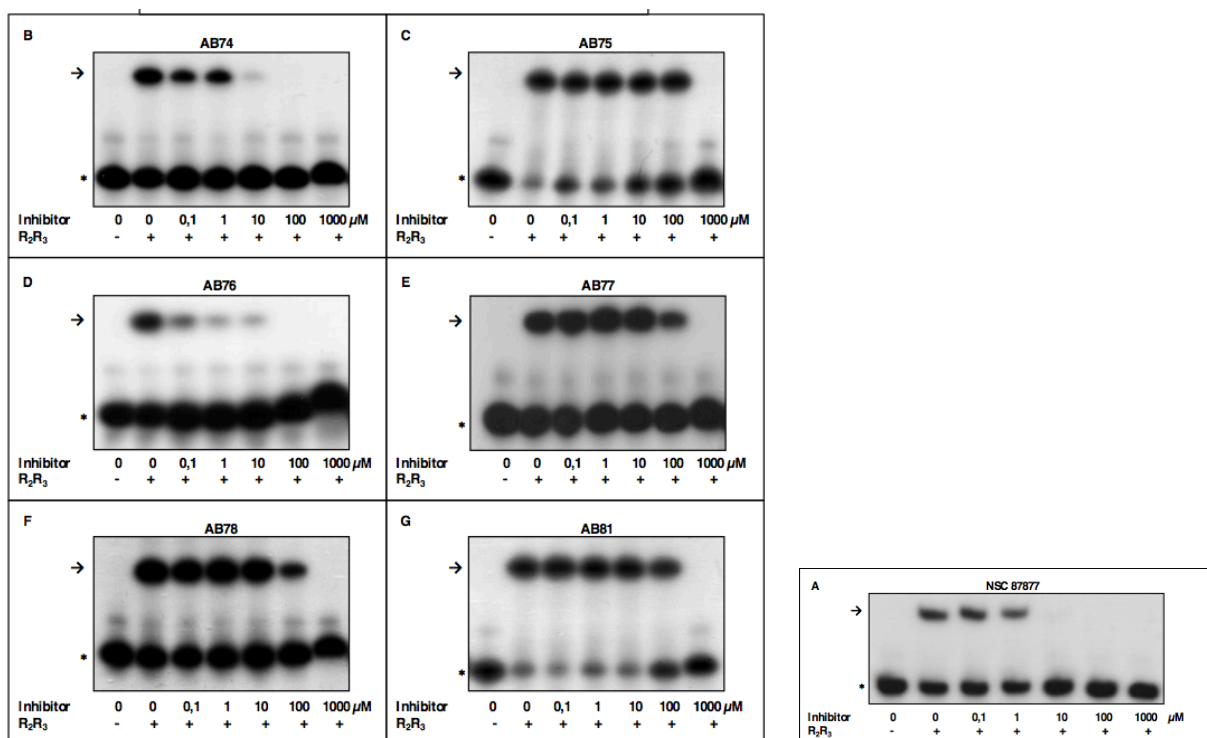


Fig. 63. Inhibition effect of AB74 - 78 and AB81. The R2/R3 regions and 0.1, 1, 10, 100 and 1000 μM of every compound was incubated with ^{32}P labeled mim-probe for 10 min. The protein/DNA complex was then separated from the unbound probe by electrophoresis. Visualization was done by autoradiography. Arrows show the protein/DNA complex, the stars show the unbound probe. The figure is taken from ref. 69.

Fig. 64. Fig. is taken from ref. 69.

An IC 50 (half maximum inhibitory concentration) value is a numerical tool that indicates the concentration of a drug molecule that is needed to inhibit a given biological process by 50 %. In our case, it is the inhibition of the binding of the R2 and R3 regions to DNA. Based on the EMSA analysis of the 6 compounds, the following IC 50 values have been found (table 8).

Compound	Free en. of bind. in R2 (kcal/mol)	R2 - Ki	Free en. of bind. in R3 (kcal/mol)	R3 - Ki	IC 50
AB 74	- 10.94	9.5 nM	- 7.23	5 μM	0.1 – 3 μM
AB 75	- 9.08	221.08 nM	- 6.49	17.35 μM	100 – 1000 μM
AB 76	- 9.32	148.38 nM	- 7.5	3.19 μM	0.1 – 10 μM
AB 77	- 9.9	55.67 nM	- 7.85	1.77 μM	50 – 100 μM
AB 78	- 9.97	48.92 nM	- 6.81	10.24 μM	100 – 200 μM
AB 81	- 9.91	53.99 nM	- 6.63	13.73 μM	100 – 1000 μM

Table 8. Comparison of the docking results with the experimental IC 50 values.

The IC 50 values differ significantly from the ones obtained in the docking studies. A closer investigation shows that the binding affinities from the docking in the R3 region correspond somewhat with the results from the EMSA-analysis. AB74, 76 and 77 have the best free energies of binding, however only AB74 and AB76 showed similar low IC 50 values. As it was shown before (figure 34), AutoDock predicted that all the top conformations of the 6 compounds that were tested would bind in the same area in the R3 region. On the contrary, in the R2 region, compounds AB74 and AB76 had the predicted mode of binding in a different region than the other four (figure 34). It is unsure whether this observation actually has something to do with the experimental IC 50 values.

All the six compounds contain two sulfonic acid groups that are capable of making hydrogen bonds with molecules that have a hydrogen donor or acceptor sites. When the compounds were synthesized, ¹H NMR analysis showed the presence of a large amount of water. One way to explain this observation is that the compounds upon precipitation co-crystallize with water molecules. Techniques such as freeze-drying did not remove water from the solid material. Most probably it is the sulfonic acid groups that are responsible for binding of water molecules. Keeping that assumption in mind, the process of the binding of the ligand to the target molecule can be viewed as follows; In order to bind to the target site, water molecules surrounding both the substrate and the drug must to some extent be expelled from the surface between the drug and the target molecule in order for interactions between the functional groups to take place - hydrogen bonds between the sulfonic acid group and e.g. the residue Lys128 must be formed. The sulfonic acid groups are very hygroscopic and make hydrogen bonds with water molecules. Theoretically, water molecules surrounding the sulfonic acid groups might not be expelled when the molecule approaches the R2 and R3 regions in c-Myb, resulting in the absence of the interactions between the drug and the residues in the target molecule. In terms of entropy, making hydrogen bonds with water molecules gives a larger positive change in entropy than binding to the target site, thus making interactions with the target site not sufficiently favorable to occur. Whether this is indeed the case, is not sure.

Another interesting fact is the results from the EMSA analysis of compound AB81. Compound AB81 was synthesized, characterized and checked by HPLC for purity, which indicated that the material was 92 % pure. Compound AB81 is the same compound as NSC87877. A small sample of NSC87877 was purchased and tested earlier by Martinez et al.^{68,69} In an EMSA analysis, the NSC87877 that was purchased gave an IC 50 value of 1 μ M (p. 54, figure 64), while the “synthesized version of NSC87877” – AB81 – had an IC 50 between 100 and 1000 μ M. The purchased sample of NSC87877 was checked by HPLC using the same conditions as used for controlling the purity of AB81. This analysis showed that the purchased sample of NSC87877 contained substances other than the desired compound – making it impossible to say what actually caused the inhibition of the R2 and R3 regions in c-Myb at the 1 μ M level.

Chapter 9

Synthesis of compounds AB 81, 74, 75, 76, 77 and 78

As was mentioned before, all of the compounds listed in figure 33 belong to a group of chemical compounds called azo-dyes. After conducting a literature search, it was decided that an azo-coupling would eventually lead to the desired products in two synthetic steps.

9.1 Theoretical aspects behind azo-coupling reaction

The material in this section is based on previously published work.⁴⁰⁻⁴³

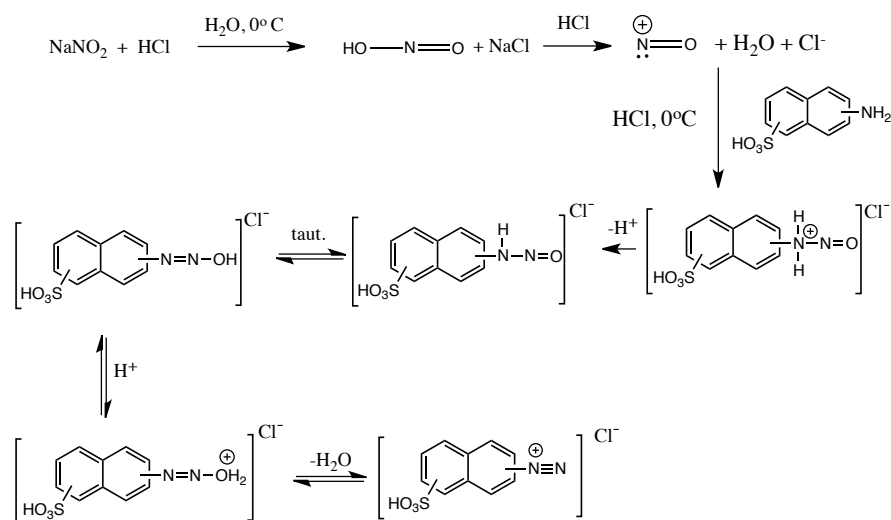
Azo-dyes are compounds containing azo-groups that are linked to sp^2 -hybridized carbon atoms. Usually the azo-groups are bound to benzene or naphthalene rings, but in some cases they can also be attached to aromatic heterocycles or enolizable aliphatic groups. Diazotization was first reported by Griess in 1858. The first azo-dyes were manufactured by Mene in 1861 and by Matius in 1863. The first dyes produced by diazotization and azo-coupling in two distinct steps were performed in 1875 by Caro and Witt.⁴⁰⁻⁴³

The diazotization of an aromatic or heteroaromatic primary amine is the first of the two reactions steps by which practically all of the azo-dyes are produced. Diazotization is usually carried out by the action of sodium nitrite on an aqueous solution of an amine in the presence of a mineral acid at 0°C (scheme 6). Since the reaction is performed in an aqueous solution, the reactants or the products exist fully or partly in their ionized forms: the acid, the nitrite and the diazonium salt exist as H^+X^- (aq), Na^+NO_2^- (aq) and Na^+X^- (aq), while the amine is in equilibrium with the corresponding ammonium ion $\text{Ar}-\text{NH}_3^+$.

According to the reaction scheme 6, at least 2 equivalents of mineral acid must be used since the pH of the solution during the diazotization must be below 2. This is necessary for several reasons:

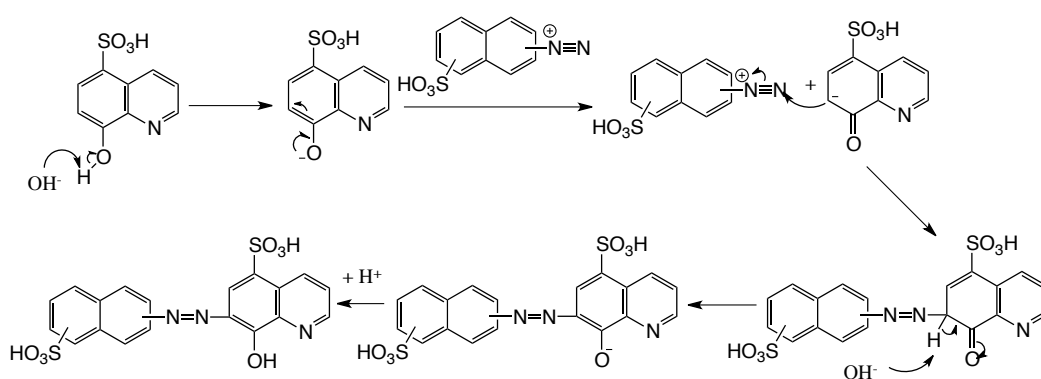
- 1) At higher pH, the equilibrium $\text{Ar}-\text{NH}_2 \leftrightarrow \text{Ar}-\text{NH}_3^+$ is shifted in the direction of the free base, which is usually less soluble in aqueous solution.
- 2) At low concentrations of H^+ ions, the diazonium ion formed reacts non-productively with the free base of an unattacked amine: $\text{Ar}-\text{N}_2\text{X} + \text{Ar}-\text{NH}_2 \rightarrow \text{Ar}-\text{N}=\text{N}-\text{NH}-\text{Ar}$
- 3) The reactive forms of the diazotizing agent are converted into ineffective ones such as HNO_2 and NO_2^- .

Excess nitrite must also be avoided, since it would exert an unfavorable effect on the stability of the diazonium ions.



Scheme 6. Generation of the diazonium salt.

Diazonium ions are Lewis acids in which the β -nitrogen atom is the center of their electrophilic character. The addition of nucleophiles at the β -nitrogen is called an azo- (or diazo-) coupling reaction and depending on the atom that provides the lone pair of electrons, C-, N-, O-, P- or S-coupling can occur. In our case the coupling partner is 8-hydroxyquinoline-5-sulfonic acid, and the coupling reaction is performed under basic conditions in order to bring sufficient nucleophilic character to the carbon atom in the *ortho*-position to the hydroxyl group (scheme 7). The sulfonic acid group occupies the *para*-position, so the *ortho*-position is the only potential position where the coupling can take place.

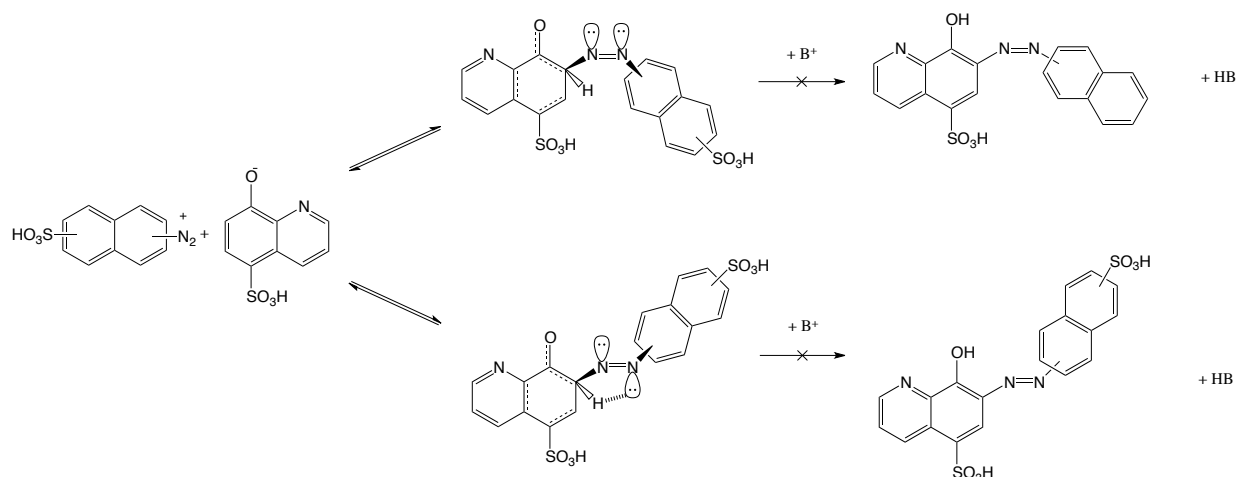


Scheme 7. The mechanism of the azo-coupling step.

It is likely that in the azo-coupling reactions the diazonium ion is added to the nucleophilic center of the substrate in such way that the σ -complex has a *trans*-configuration (scheme 8). Although there is no strict evidence against a primary formation of a compound with a *cis*-configuration, Zollinger showed

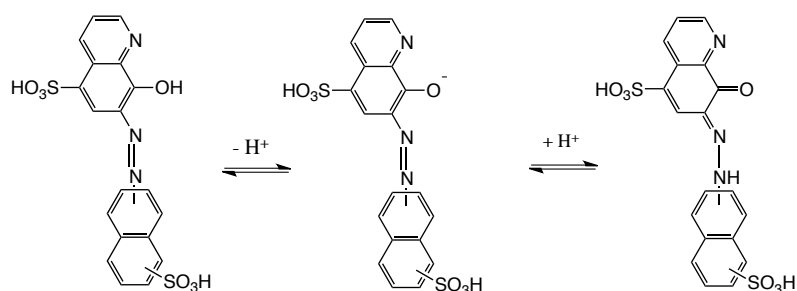
that all additions of “hard” nucleophiles to the β -nitrogen of a diazonium ion (e.g. OH^-) have an early transition state and the *cis*-configuration is the primary product. With “soft” nucleophiles (C- and N-coupling) a later transition state is feasible; *trans*-isomers are formed. The *cis*-isomers are available only by photochemical isomerisation of the *trans*-compounds.

Another argument can be drawn from the steric considerations of the transition state leading to the formation of the desired diazo-compound. As it can be seen in scheme 8, the transition state leading to the formation of the *cis*-product do not provide enough space for the base to come in and eliminate the hydrogen atom which is *ortho* to the hydroxyl/keto group. The transition state leading to the formation of the *trans*-product does provide enough space for the base to come in and abstract the proton. An additional argument can be drawn from the possible stabilization of the transition state leading to the *trans*-product by electrostatic forces between the hydrogen atom *ortho* to the hydroxyl/keto group and the lone pair of electrons on the α -nitrogen.



Scheme 8. Representation of the two possible transition states leading to the formation of the diazo-compounds.

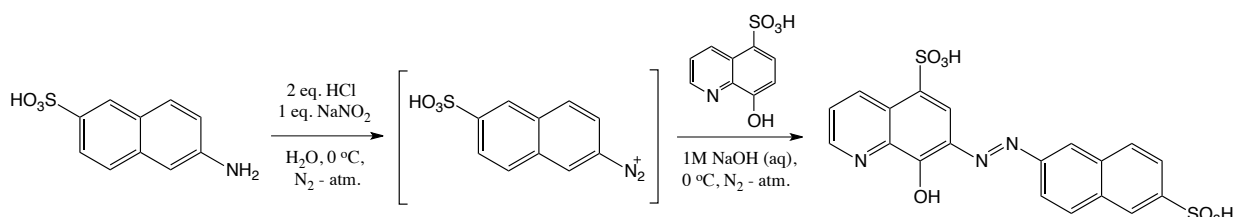
The diazo-compounds in solution exist most likely as different tautomeres, as it is shown (scheme 9). Which of the forms that is most favorable, depends on the pH of the solution.



Scheme 9. Tautomerization of the diazo-compounds.

9.2 Synthesis, results and discussion

Synthesis of the series of the compounds AB81 - 78 began with the preparation of compound AB81 that has the IUPAC name of 8-oxo-7- [(6-sulfonaphthalen-2-yl)hydrazinylidene]quinoline-5-sulfonic acid. As it can be seen from the name, the two coupling partners for the azo-coupling reaction were the 8-hydroxyquinoline-5-sulfonic acid and the 6-aminonaphthalene-2-sulfonic acid. The overall reaction procedure can be viewed in scheme 10.



Scheme 10. The overall reaction procedure leading to the formation of the diazo-product.

From the literature it was known that azo-dyes containing the sulfonic acid groups could be precipitated from a crude reaction mixture at a pH of 10 - 12 by addition of concentrated hydrochloric acid.⁴⁰⁻⁴³ Unfortunately, during the first attempts of the synthesis of AB81, absolutely no material could be precipitated by addition of acid (the reaction mixture was allowed to stay at room temperature for 2 hours after addition of the acid). A first thought was that the diazotation step did not go as planned – the diazonium salts are not stable and commonly decompose at temperatures above 5 or 10 °C. Upon repetition of the experiment, the solution of sodium nitrite in water was added over 10 min to the solution of amine in water that had a temperature of -5 °C. The color change from clear to yellow indicated that the diazonium salt was formed – when an aromatic system such as naphthalene is extended by the addition of a triple bond, a color change usually follows in the direction of colors that have longer wavelengths. Upon warming of the solution to room temperature, bubbling was noticed which can be explained in terms of a decomposition of the diazonium salt and release of nitrogen gas. Acting upon this, attention was focused on the second step of the reaction – the coupling process itself. Three experiments were set up to test the coupling step at pH 8, 10 and 12. Upon addition of the 8-hydroxyquinoline-5-sulfonic acid dissolved in 1M NaOH to the solution of diazonium salt at -5 °C, an intensive color change towards red-purple was observed at all three pH values. Using the same arguments about extension of the conjugated system and color change, it can be argued that the coupling reaction had occurred. However, upon addition of the concentrated hydrochloric acid, no precipitation was observed.

Further investigation of the literature uncovered a very interesting detail: prior to the diazotation, 1 eq. of sodium carbonate was added to the solution of the amine in water. If formation of the sodium salt of

the amine would somehow affect the reaction, was not clear, so an experiment was set up where sodium carbonate was added to the solution of the amine – the other experimental parameters were left unchanged. Surprisingly, precipitation of the crude product upon addition of hydrochloric acid proceeded smoothly, making it possible to filter off the crude product. Purification of the crude product was carried out by dissolving it in hot water, followed by hot filtration and precipitation of the product with dioxane. But since dioxane is a toxic solvent, it was decided to purify the crude product instead by dissolving it in 1M NaOH and then adding concentrated hydrochloric acid to reprecipitate the product. An ^1H NMR showed the absence of signals from the starting materials after this purification protocol was carried out 2 - 3 times.

The synthesis of the other 5 compounds followed the same reaction protocol as for AB81. Some amines were not soluble in water at room temperature and required heating to 60 or 70 °C in order to dissolve. Precipitation of the desired products with concentrated hydrochloric acid commonly took from 5 minutes to 1 hour, but in some cases it took up to 12 hours.

The yields of the azo-coupling reaction were very low – between 15 and 35%. It is not clear whether this is caused by side reactions or decomposition of the diazonium salt. Also the recoveries from the reprecipitation technique were low, so it is plausible that some percentage of the desired product was lost during purification.

An investigation of the crude and pure products by TLC technique proved to be very challenging. TLC analysis on the normal phase silica plates using water/acetic acid (1:1) gave a very strong tailing. Eluent systems as methanol/water, methanol/acetic acid, acetonitrile/water also did not work. When TLC analysis was performed using reverse phase silica plates in acetonitrile/acetic acid (5:2), less tailing was observed than in the previous cases, but it was not enough for a clear distinction of the different spots.

The purity of the synthesized azo-compounds was checked using a reverse phase HPLC system with acetonitrile/ammonium acetate buffer (pH 10) as eluent system. HPLC conditions were obtained from TOCRIS Bioscience.⁷⁰

NMR characterization of the synthesized compounds included ^1H NMR spectra of the compounds purified both by precipitation with dioxane and hydrochloric acid. In the ^1H NMR spectra a sharp peak at 3.55 ppm and in ^{13}C NMR a peak at 66 ppm corresponded to dioxane. When the purification was done using only 1M NaOH and concentrated hydrochloric acid, the sharp peak at 3.55 ppm disappeared. A very broad peak around 4 ppm corresponded to water present in the sample. It was attempted to remove water by freeze drying techniques, but even after 3 days on a freeze dryer, water was still present in the samples in high concentrations.

Chapter 10

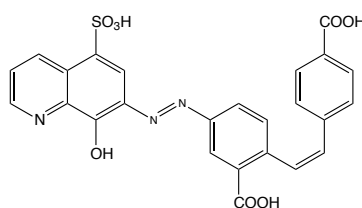
Synthesis of the compound NSC675460

10.1 Introduction

The target compound NSC675460 was identified as a potential inhibitor of the DBD in c-MYB after screening of the NCI library for substances that exhibit structural and electronic similarities with NSC87877.⁶⁸ After a literature research, it became clear that the compound has never been synthesized before and exists only as a “virtual” compound in the NCI library. Therefore, it was decided to explore different strategies to synthesize NSC675460.

10.2 Overview of NSC675460

NSC675460 consists of two aromatic phenyl rings, both bearing carboxylic acid groups. The phenyl rings are connected by a C=C *cis*-double bond and one of the rings is connected to a functionalized quinoline moiety by an azo-bridge. One of the first concerns was the stability of the *cis*-double bond. The possibility of spontaneous isomerisation to the more stable *trans*-analogue had to be taken into account by avoiding chemicals and conditions that could facilitate isomerisation. Another aspect was the polarity of the final product, which could make purification by standard methods like flash chromatography quite challenging.

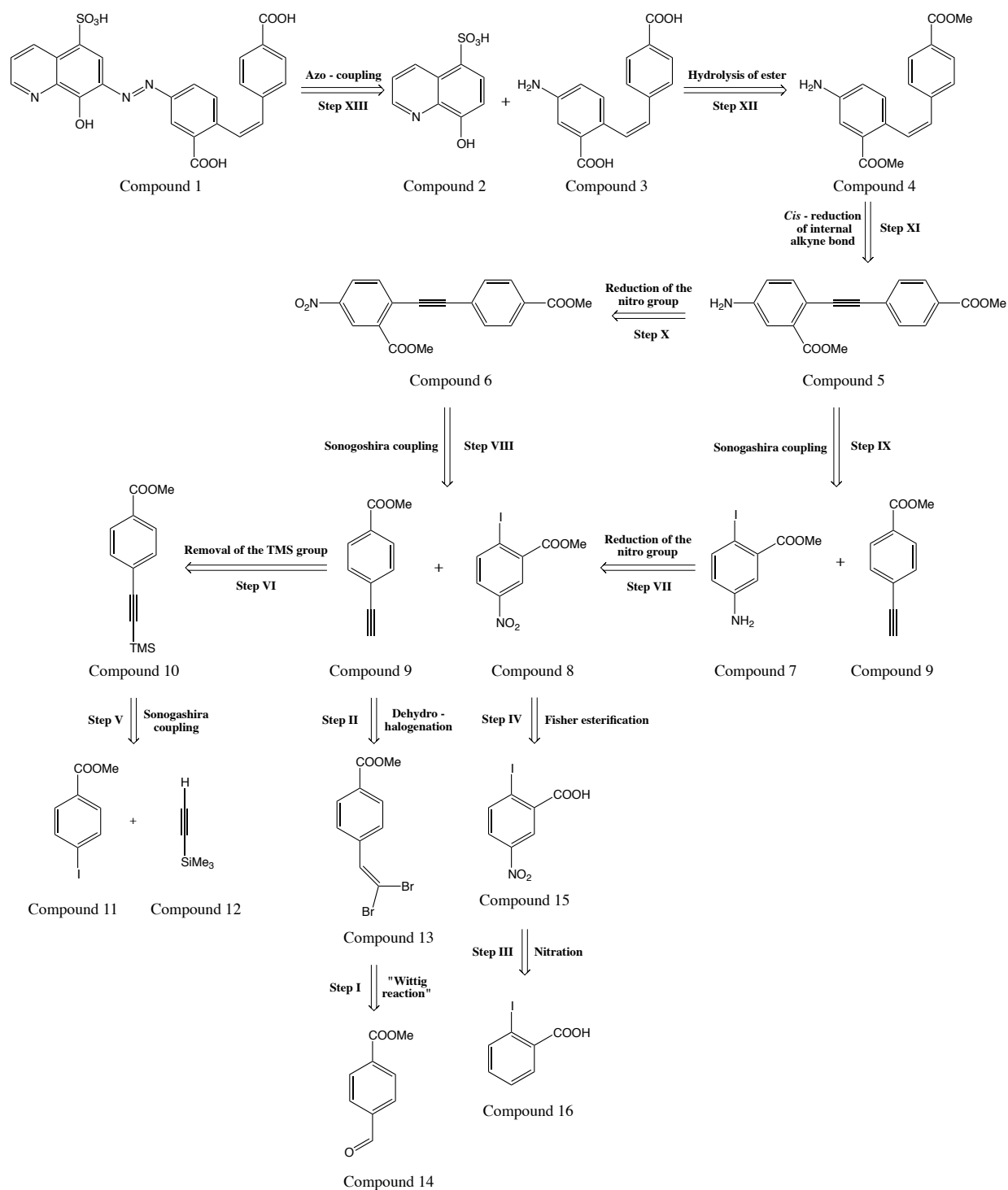


NSC675460

10.3 The retrosynthetic analysis

The retrosynthetic analysis began by considering possible disconnections in the target compound (scheme 11). It became clear that a disconnection of the quinoline moiety from the rest (step XIII) would simplify the synthesis to a great extent. The availability of the 8-hydroxyquinoline-5-sulfonic acid from commercial sources and the fact that the azo-bridge can be made in a two-step reaction avoiding harsh reactions conditions were the main thoughts behind the disconnection. This disconnection gave the next retrosynthetic intermediate with the IUPAC-name (*Z*)-5-amino-2-(4-

carboxystyryl)benzoic acid (compound 3, scheme 11). After an extensive literature search it became clear that the compound had not been previously reported. The compound contains an internal *cis*-double bond – one of the key features in the target molecule of our interest. It was decided that the replacement of the *cis*-double bond with an internal triple bond would provide a more synthetically robust intermediate.



Scheme 11. The retrosynthetic analysis of NSC675460.

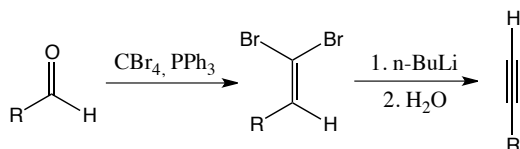
The two carboxylic acid groups present in the target molecule could readily be derived from the methyl- or ethyl esters, which are easier to work with than carboxylic acids with respect to purification by flash chromatography and identification of synthetic intermediates by NMR- and TLC-analyses – esters provide easily recognizable peaks in ¹H NMR spectra and do not give tailing on TLC plates when using common eluent systems. This brings us to the compound 5 in scheme 11, with the IUPAC-name methyl 5-amino-2-((4-(methoxycarbonyl)phenyl)ethynyl)benzoate. A literature search showed that this compound had not been reported before. Thus, the next step was the analysis of possible methods for constructing an internal alkyne bond – the Sonogashira coupling reaction became our method of choice since it often provides desired alkyne product in good yields, does not require harsh reaction conditions and also tolerates the presence of different functional groups. This retrosynthetic step gave two intermediates; methyl 5-amino-2-iodobenzoate and methyl 4-ethynylbenzoate. The former of which has not been prepared before while the latter is a known compound. The methyl 5-amino-2-iodobenzoate could be prepared from the 2-iodobenzoic acid by nitration, followed by methylation of the carboxylic acid and reduction of the nitro group, as shown in the scheme 11. The synthesis of the methyl 4-ethynylbenzoate by the original route shown in the scheme 12 proved to be very problematic, and thus the compound was later synthesized by another procedure.

The retrosynthetic analysis was carried out before the experimental work was started. As such, parts of the analysis and the order of some of the steps proved to be wrong or misleading, as it always happens. The synthesis of the methyl 4-ethynylbenzoate by the original route proved to be difficult and it was thus prepared by another procedure. The order of some of the synthetic steps was also changed in order to obtain a more convergent synthetic pathway leading to the desired compound.

10.4 Syntheses, results and discussion

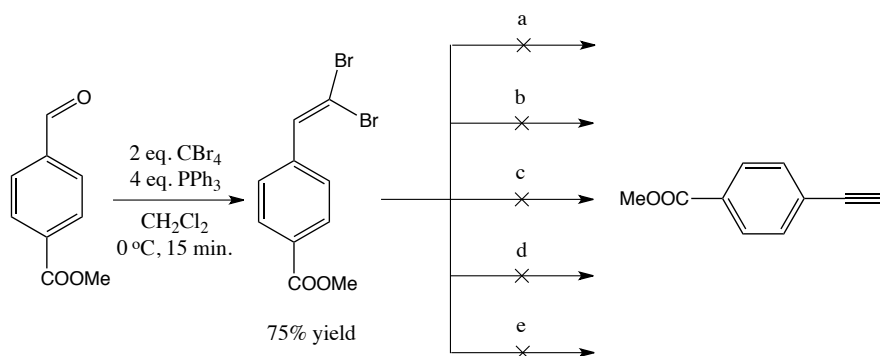
10.4.1 Corey-Fuchs reaction

The synthetic work began with the synthesis of the methyl 4-ethynylbenzoate (compound 9, scheme 11). The first idea was to synthesize it from methyl 4-formylbenzoate by performing Corey–Fuchs reaction.⁷¹ The Corey–Fuchs reaction is a two-step reaction where an aldehyde is transformed into a terminal alkyne by a one-carbon homologation. The first step of the reaction can be compared to the Wittig reaction and leads to a dibromoolefin (scheme 12). Treatment of the dibromoolefin with a lithium base, in the second step, leads via dehydrohalogenation to the bromoalkyne intermediate, which then undergoes a metal-halogen exchange which yields the terminal alkyne upon aqueous work-up (scheme 12).⁴⁴



Scheme 12. Corey – Fuchs reaction

The first step in the reaction proceeded smoothly and gave the desired product in good yields (scheme 13). The second step of the reaction – the dehydrohalogenation of the dibromoalkene with subsequent metal-halogen exchange and generation of the terminal alkyne – did not work. Several techniques and reagents were tested (scheme 13). None of them gave the desired product. The dehydrohalogenation with different amounts of *n*-BuLi at different temperatures, followed by an aqueous work-up gave product mixtures that could not be identified or purified by flash chromatography. ¹H NMR of the crude product mixture showed the absence of the methyl ester signals. This suggested that *n*-BuLi, even at low temperatures had attacked the electrophilic methyl ester group in favor of abstracting the weakly acidic proton on the dibromoalkene. Another experimental procedure was tested in order to see if it would give the terminal alkyne. It is possible to perform the second step of the reaction – the dehydrohalogenation – by refluxing the dibromoalkene in THF at 120 °C in the presence of activated magnesium metal.⁷² But despite all the efforts, a ¹H NMR-analysis of the crude mixture indicated that no reaction took place.



a) Mg, activated with I_2 , THF, 120 °C, 2 hours; b) Mg, activated with $C_2H_2Br_2$, THF, 120 °C, 2 hours; c) Mg, activated by washing with 1 M HCl, then heating under vacuum, THF, 120 °C, 2 hours; d) 2 eq. *n*-BuLi, -78 °C → 0 °C → r.t., H_2O , 4 h.; e) 3 eq. *n*-BuLi, -78 °C, H_2O , 4 h.

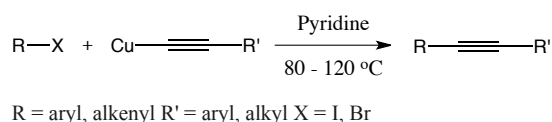
Scheme 13. Attempted Corey – Fuchs reaction.

As a result of this failure, it was decided to synthesize the methyl 4-ethynylbenzoate in two steps. The first step was the Sonogashira coupling of the methyl 4-iodobenzoate with the ethynyltrimethylsilane, giving the methyl 4-((trimethylsilyl)ethynyl)benzoate.

10.4.2 The Sonogashira coupling reaction

The palladium catalyzed C-C homologation reaction that couples terminal sp -hybridized carbon atom of an alkyne with an sp^2 hybridized atom of an aryl or vinyl halide is typically referred to as the Sonogashira coupling reaction. There are many variations of this reaction, but the most general synthetic protocol uses copper as a co-catalyst, in addition to the palladium catalyst. During the last two decades, the Sonogashira reaction has seen an incredible growth in its versatility, proving its role as an important tool in organic synthesis. There are many examples of applications of the Sonogashira coupling in syntheses of natural products, biologically active molecules and technologically advanced materials. An arylalkyne may provide a molecule with a rigid structure and a specific shape, but it also gives an unsaturated, high-energy moiety that can undergo further synthetic transformations.

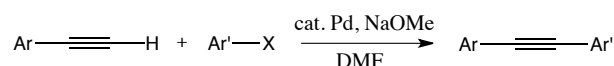
The first example of an sp^2 - sp cross-coupling between two carbon atoms was reported in 1963 by Stephens and Castro.⁴⁵ A general overview of the reaction protocol can be seen in scheme 14.



Scheme 14. Stephens and Castro's coupling method.

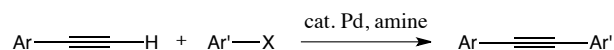
During the following years, a number of other synthetic protocols appeared dealing with the cross-coupling of sp^2 - sp hybridized carbon atoms (schemes 15 - 17).

Cassar's method:⁴⁶



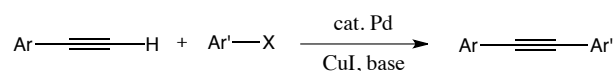
Scheme 15. Cassar's coupling method.

Heck's method:⁴⁷



Scheme 16. Heck's coupling method.

Sonogashira's method:⁴⁸

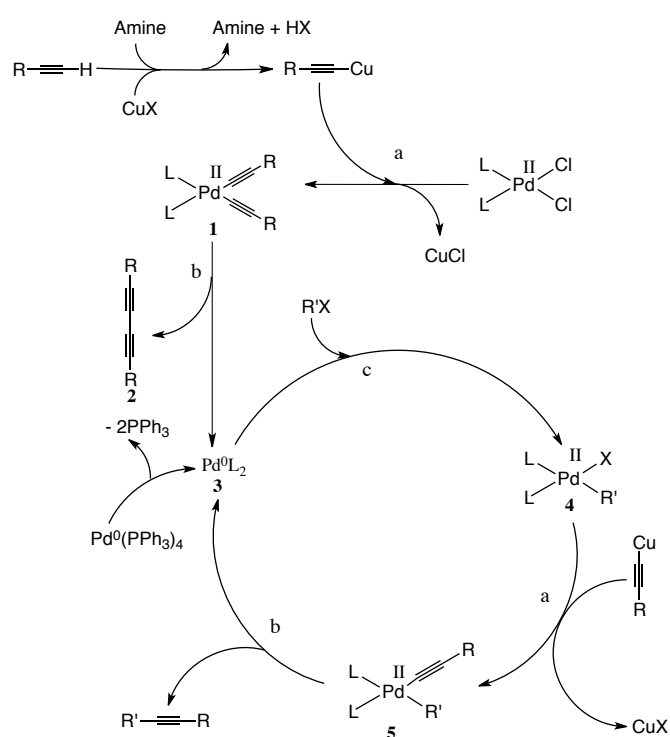


Scheme 17. Sonogashira's coupling method.

Sonogashira's protocol may be seen as a combination of three other protocols. As a matter of fact, the catalytical amount of CuI increased the reaction yield and made it possible to conduct reactions at lower temperatures. One of the main differences between the Sonogashira reaction and other C–C homologation reactions involving a palladium catalyst, is that the alkyne does not need to be activated by a halide or a metal. The alkyne itself serves as a good ligand for the palladium catalyst because of the shape of its *d*-orbitals and the absence of steric bulk. It has also been shown that the Sonogashira reaction tolerates nearly all common functional groups as a consequence of the mild reaction conditions.⁴⁹ Such a wide functional group tolerance makes it an attractive procedure for the construction of alkyne moieties in syntheses of complex molecules.

Mechanism

The accepted mechanism of the Sonogashira reaction does not differ that much from the original one proposed by Sonogashira and Hagira (scheme 18).⁵⁰ However, there are some evidence that suggests a more complex reaction mechanism.^{50a, 51}



a: Transmetalation; b: Reductive elimination; c: Oxidative addition

Scheme 18. The accepted mechanism of the Sonogashira coupling reaction

The most commonly used catalyst for the reaction is either Pd(0)(PPh₃)₄ or Pd(II)(PPh₃)₂Cl₂. The catalyst loading is typically 2 – 5 mol%, the amount of the CuI is usually twice as high. There are also

reported reactions where the loading of the palladium catalyst is lower and the reaction proceeds without a copper co-catalyst. Demanding substrates may require a different catalyst system that is electronically and sterically tuned to the specific reaction.

In the Sonogashira coupling reaction, the catalytically active complex is Pd(0)L₂ (**3**, scheme 18). Following the classical reaction mechanism, this species is generated from Pd(0)(PPh₃)₄ by dissociation of the two phosphine ligands. It can be also generated from Pd(II)(PPh₃)₂Cl₂ via transmetallation of an alkynyl copper adduct, followed by reductive elimination of (**1**) in scheme 18 to (**3**) and (**2**). The amount of the by-product (**2**) can be minimized by removing oxygen from the reaction mixture by purging the reaction mixture with an inert gas before and during the slow addition of the alkyne coupling partner. Running the reaction under inert atmosphere also diminishes amount of the by-product (**2**). After the formation of complex (**3**), oxidative addition of the aryl or vinyl halide occurs giving complex (**4**). Transmetallation of complex (**4**) with the alkynyl copper adduct gives (**5**), followed by reductive elimination that furnishes the cross-coupled product and regenerates the catalytically active complex (**3**).

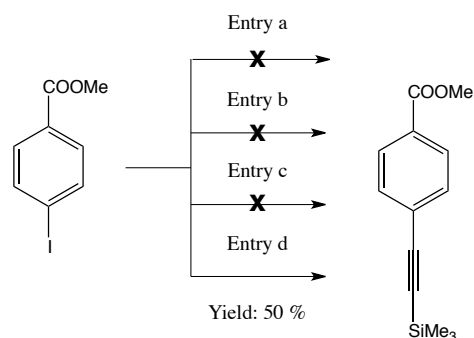
There are many factors that govern the efficiency of the catalytic system. These factors usually evolve around the ligands on the palladium-center, the amine base, the copper co-catalyst, the solvent effects and the steric and electronic environments of the aryl/vinyl halide and the alkyne. Organohalides containing electron-withdrawing groups are commonly more reactive in the Sonogashira coupling reaction than ones with electron-donating groups. For alkynes the opposite is true.⁵²

As it can be seen from scheme 18, the amine base is an important component of the Sonogashira reaction. Et₃N, Et₂NH and *i*Pr₂NH are commonly used bases and often give good results. It must, however, be pointed out that the choice of base is sometimes substrate dependent and is usually a result of trial and error. Stronger bases may give an increase both in the reaction rate and the yield according to some publications.⁵³ In the classical Sonogashira protocol, the amine acted both as the base and the solvent. Later research indicated that use of co-solvent – mostly THF – gave increases in the reaction rates and yields. It is difficult to give a detailed explanation of this effect – substrate and catalyst solubility might be a major factor.

10.4.3 The Sonogashira coupling of the methyl 4-iodobenzoate with the ethynyltrimethylsilane

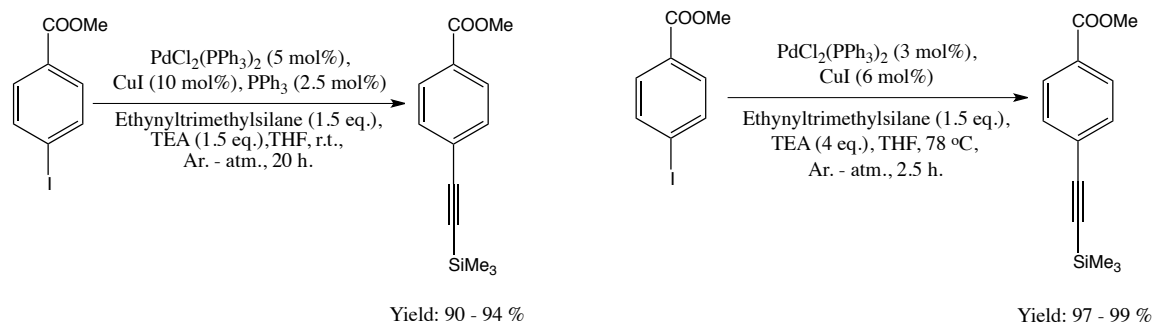
The first attempts on the synthesis of methyl 4-((trimethylsilyl)ethynyl)benzoate were performed using the “classical” protocol in which the active catalyst – Pd(0)(PPh₃)₄ – was generated *in situ* from

$\text{Pd}(\text{OAc})_2$ and PPh_3 . Et_3N was used both as a solvent and a reagent (scheme 19, entry b). ^1H NMR analysis of the crude reaction mixture did not show any conversion of the starting material. The use of DMF as a solvent and the addition of a smaller amount of Et_3N also did not give any conversion (scheme 19, entry a). After the first unsuccessful attempts, the attention was moved to another catalyst which is commonly used in the Sonogashira reaction – $\text{Pd}(\text{II})\text{Cl}_2(\text{PPh}_3)_2$. The first attempt (scheme 19, entry c) was unsuccessful, but the second one gave the desired product in 50 % yield (scheme 19, entry d). After further investigation of the reaction protocols, two slightly different approaches were employed for the synthesis of methyl 4-((trimethylsilyl)ethynyl)benzoate, giving the desired product in good to excellent yields (schemes 20 and 21). The reaction protocol shown in scheme 21 provided the product in less time than the other methods, with a lower loading of the palladium catalyst. The procedure was reproducible and as such it was chosen for the synthesis of methyl 4-((trimethylsilyl)ethynyl)benzoate.



Entries: a) $\text{Pd}(\text{OAc})_2$ (5 mol%), CuI (10 mol%), PPh_3 (20 mol%), TMSA (1 eq.), TEA (0.7 eq.), DMF , r.t., N_2 - atm., 16 h.;
 b) $\text{Pd}(\text{OAc})_2$ (5 mol%), CuI (10 mol%), PPh_3 (20 mol%), TMSA (1 eq.), TEA (10 eq.), 80 °C, N_2 - atm., 6 h.;
 c) $\text{PdCl}_2(\text{PPh}_3)_2$ (2 mol%), CuI (1 mol%), TMSA (1.2 eq.), TEA (5 eq.), r.t., N_2 - atm., 20 h.;
 d) $\text{PdCl}_2(\text{PPh}_3)_2$ (3 mol%), CuI (6 mol%), TMSA (1.3 eq.), TEA (7 eq.), r.t., N_2 - atm., 20 h.

Scheme 19. The attempted Sonogashira coupling reactions between the methyl 4-iodobenzoate and the TMSA.



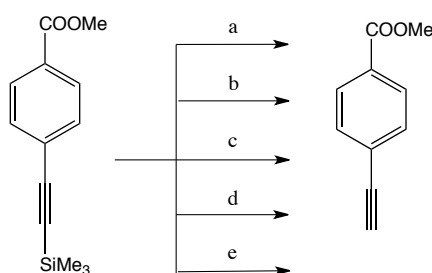
Scheme 20. Modified protocol of the Sonogashira coupling reaction.

Scheme 21. Optimized protocol of the Sonogashira coupling reaction.

There are several factors that might explain the failed reactions in scheme 19. Generation of the $\text{Pd}(\text{PPh}_3)_4$ *in situ* might be problematic – any air present in the solvent or reaction flask will prevent the formation of the catalyst or destroy one already formed. $\text{PdCl}_2(\text{PPh}_3)_2$ on another hand is an air stable catalyst and can be purchased from commercial sources. Employing THF as the solvent and adding the required amount of the amine (1.5 – 4 eq.), instead of running the reactions in neat Et_3N might also have contributed to the increase in the yields. Careful deoxygenation of the solid reactants, purging of the solvents and liquid reactants with argon before mixing was performed in order to provide an inert atmosphere during the reaction. The reaction itself was done under continuous flow of argon.

10.4.4 Removal of the TMS group

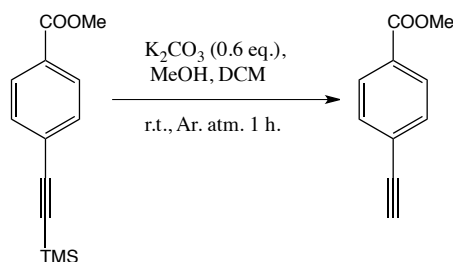
The next step was the removal of the TMS group from the methyl 4-((trimethylsilyl)ethynyl)benzoate. There are many procedures and techniques for carrying out this step (scheme 22).⁷⁶



a) K_2CO_3 , MeOH, r.t.;^{76a} b) DBU, H_2O , MeCN, 60°C ;^{76b} c) K_2CO_3 , MeOH, CH_2Cl_2 , r.t.;^{76c} d) CsF, MeCN, H_2O , r.t.;^{76d} e) *t*- Bu_4NF , THF, r.t.^{76e}

Scheme 22. Methods dealing with removal of the TMS group based on the literature.

K_2CO_3 in MeOH usually cleaves the TMS group selectively and in good yields. With respect to the elimination of the TMS group, some attention must be addressed to the amount of K_2CO_3 present in the reaction mixture. Experiments showed that when a large excess of K_2CO_3 is used, hydrolysis of the methyl ester might occur, especially when the reaction is performed over prolonged periods of time. It is also known that KF and TBAF readily cleaves carbon-silicon bonds and do not affect base labile groups, but there are some serious drawbacks when it comes to handling KF and TBAF. In addition to their toxicity, their hygroscopic nature also provides some additional difficulties since it is not a trivial task to measure amount of water present thus making it difficult to weight out precise amounts of the reagents. After some experimental work and a literature search, a procedure for a fast and reliable removal of the TMS group was set up. K_2CO_3 (0.6 eq.) in MeOH/ CH_2Cl_2 (3:1) furnished the desired product after 1 hour at room temperature (scheme 23).



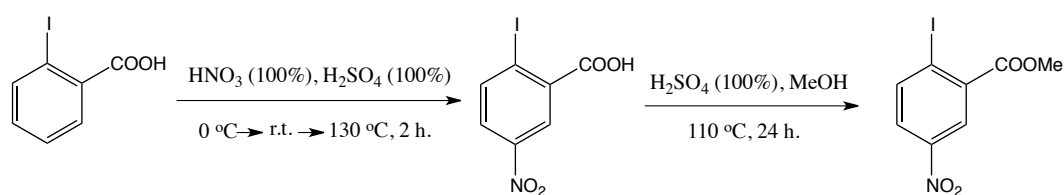
Scheme 23. Procedure for fast and reliable removal of the TMS-group.

The yields were very good: 95 – 99%. Pure methyl 4-ethynylbenzoate is a white solid, but in some cases it had either yellow or reddish color. Upon storage at 0 °C overnight under an argon atmosphere, a black substance appeared inside the flask in most cases. It was easily removed by flash chromatography, resulting in a 5 – 8% loss of the compound. In most cases, the methyl 4-ethynylbenzoate was used directly in the next reaction.

With one of the major building blocks synthesized, the focus was changed towards the preparation of another building block – the methyl 5-amino-2-iodobenzoate. At this point there were some doubts regarding how well the Sonogashira coupling of the methyl 4-ethynylbenzoate with the methyl 5-amino-2-iodo-benzoate would work since the amino group is a strongly electron-donating group. It was thus decided to synthesize the methyl 2-iodo-5-nitrobenzoate instead, couple it with the methyl 4-ethynylbenzoate and then reduce the nitro-group.

10.4.5 Preparation of the methyl 2-iodo-5-nitrobenzoate

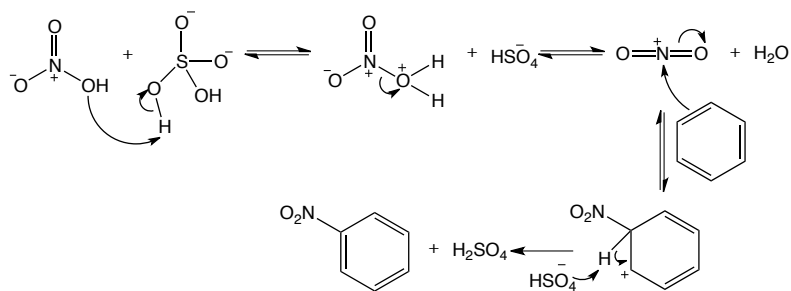
Methyl 2-iodo-5-nitrobenzoate has been prepared in two steps as shown in scheme 24. The first step is the nitration of 2-iodobenzoic acid and the second step is the esterification of the carboxylic acid group.



Yield over 2 steps: ca 60 %

Scheme 24. A general protocol for the nitration and esterification of the 2-iodobenzoic acid.

Nitration of the 2-iodobenzoic acid is a classical example of electrophilic aromatic substitution (scheme 25).



Scheme 25. The mechanism of nitration.

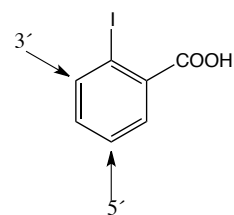


Fig. 65. Potential positions of nitration.

Both the iodine and the carboxylic acid group are deactivating substituents, with the iodine being an ortho-, para-directing group and the carboxylic acid being meta-directing functional group. Taking these considerations into account, the 2-iodobenzoic acid could in theory be nitrated both in the 3'- and 5'-position (figure 65). Since the iodine atom and the nitro group are quite bulky substituents, nitration in the 3'-position was considered highly unrealistic and thus the nitrated product would have the nitro group in the 5'-position. As was mentioned earlier, the iodine and carboxylic acid are deactivating substituents, so high reaction temperatures and long reaction times were required for the formation of the 2-iodo-5-nitrobenzoic acid. This reaction was performed by mixing the starting material with concentrated H_2SO_4 at 0°C , adding cold 99% HNO_3 mixed with concentrated H_2SO_4 and then stirring the reaction mixture at 130°C for 2 hours. This furnished the desired product in 75 – 85% yield. Experiments where 66% HNO_3 was used did not furnish the desired product.

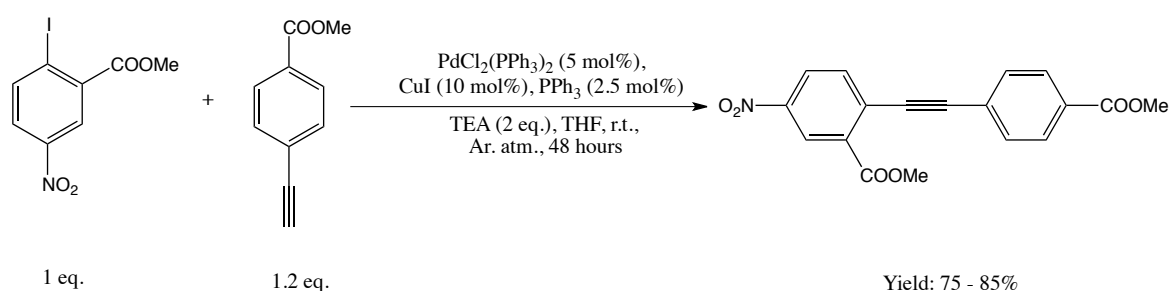
Esterification of the carboxylic acid group was performed using a classical method – the Fischer esterification. Refluxation of the crude 2-iodo-5-nitrocarboxylic acid in the mixture of the H_2SO_4 and MeOH at 110°C for 18 hours. The yields were around 75 – 85%. Separation of the crude 2-iodo-5-nitrobenzoic acid was carried out by pouring the cold reaction mixture over ice/water and then filtering the solution on a Büchner-funnel. The crude product contained large amounts of water that could affect the next step – the esterification in methanol with concentrated H_2SO_4 . According to most reaction protocols, the compound to be esterified had to be dry since water quenches the catalytic activity of H_2SO_4 . However, after drying the crude 2-iodo-5-nitrobenzoic acid with either phosphorus pentoxide overnight or under open air for 2-3 days, the esterification did not proceed. This indicated that the 2-iodo-5-nitrocarboxylic acid was unstable and therefore the crude product had to be used in a wet state in the esterification step. In the “classical” esterification protocol, only a catalytical amount of H_2SO_4 is added to the reaction mixture, while the alcohol is present in excess. In this case, methanol could not dissolve the crude 2-iodo-5-nitrobenzoic acid alone, even at 110°C . However, after the addition of a large amount of H_2SO_4 ($\text{MeOH}/\text{H}_2\text{SO}_4$, 2 – 3:1), the crude 2-iodo-5-nitrobenzoic acid finally dissolved at 110°C and the reaction furnished the desired product in 75 – 85 % yield. An interesting fact is that

the methyl 2-iodo-5-nitrobenzoate precipitates immediately upon cooling of the reaction mixture in an ice bath.

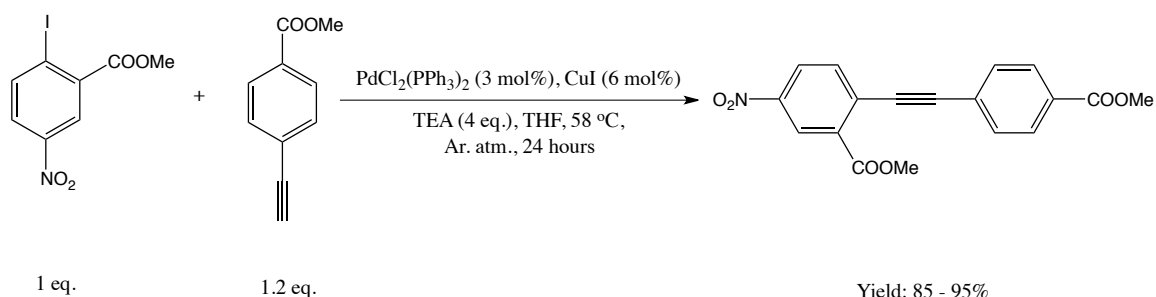
The next steps were the Sonogashira coupling reaction of the methyl 4-ethynylbenzoate both with the methyl 2-iodo-5-nitrobenzoate and the methyl 5-amino-2-iodobenzoate as well as the reduction of the nitro group to an amino group in the case of methyl 2-iodo-5-nitrobenzoate. As mentioned, most probably the reduction would be performed on the coupled product, methyl 2-(4-(methoxycarbonyl)styryl)-5-nitrobenzoate.

10.4.6 The synthesis of methyl 2-(4-(methoxycarbonyl)styryl)-5-nitrobenzoate

The Sonogashira coupling between methyl 4-ethynylbenzoate and methyl 2-iodo-5-nitrobenzoate was carried out using the same methods as the coupling between the methyl 4-iodobenzoate and the ethynyltrimethylsilane (schemes 26 and 27).



Scheme 26. Method 1. The Sonogashira coupling between the methyl 4-ethynylbenzoate and methyl 2-iodo-5-nitrobenzoate.



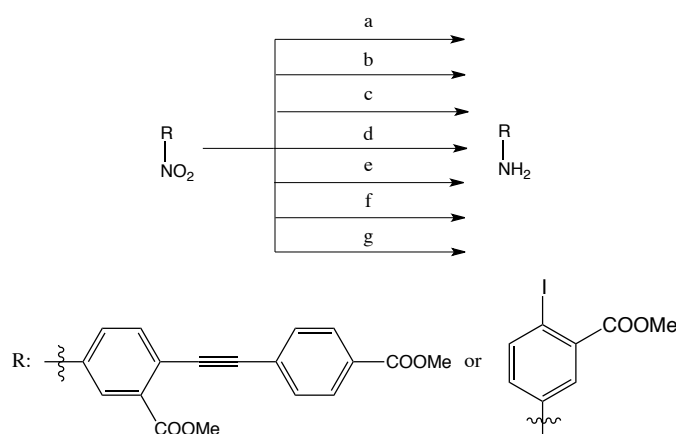
Scheme 27. Method 2. The optimized Sonogashira coupling protocol.

Use of less than 1.2 equivalents of the ethynyl coupling partner resulted in a major decrease in yield. As was pointed out before, pure methyl 4-ethynylbenzoate sometimes had reddish colour – in these cases coupling did not proceed smoothly with either of the two methods and yields were 10 – 15 % lower. Methyl 4-ethynylbenzoate is not a stable compound (section 10.4.4), an observation that may explain the lower yields caused by the longer reaction times of method 1 (scheme 26). On the other hand, the reaction temperature in method 2 is higher, making it unrealistic that the methyl 4-ethynylbenzoate

would decompose slower than the method 1. Possibly however, the rate of coupling in method 2 could have been higher than the rate of decomposition of the starting material, giving higher reaction yields. Reaction temperatures above 58 °C did not increase the yield.

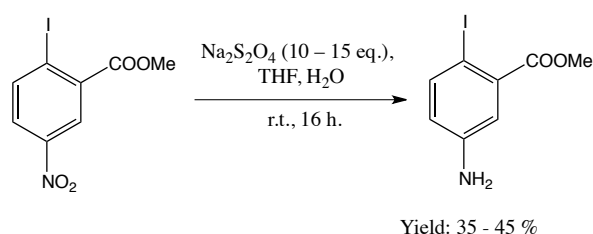
10.4.7 The reduction of the nitro group and hydrolyses of the ester groups

There are numerous reaction protocols dealing with the reduction of both aromatic and aliphatic nitro groups.⁵⁴ Some methods were chosen for further investigation and testing on our substrates of interest (scheme 28). The methods employing metals and acids seemed to be the most appropriate since protocols where a palladium catalyst and hydrogen gas was being used could potentially react with the triple bond and give unwanted reduction products. The protocol involving the use of sodium dithionite in water/THF solution was the first to be tested, since it is a known and very mild procedure for the reduction of aromatic nitro groups. An excess of the sodium dithionite (10 – 15 eq.) was used in the reduction of the nitro group on both substrates, giving the desired products in 35 – 50 % yield (schemes 29 and 30). Attention was thus moved to the other methods employing metals and acids.

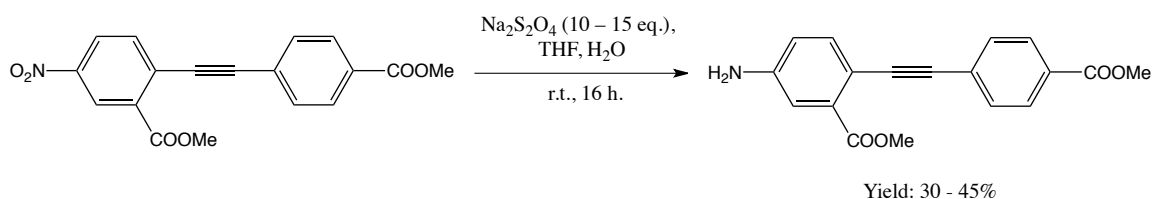


a) SnCl_2 , HCl , solv. / reag.; b) Zn , NH_4Cl (aq), H_2O , THF , r.t., 6 – 24 h.; c) Pd , CO_2 , H_2 , 180 bar, elevated temperature; d) 1. FeSO_4 , H_2O , MeOH , reflux. 2. NH_4OH , H_2O , reflux; e) Zn , NH_4Cl , MeOH , H_2O , r.t., 6 – 24 h.; f) $\text{Na}_2\text{S}_2\text{O}_4$, H_2O , THF , r.t., 8 – 24 h.; g) Fe , CaCl_2 , EtOH , H_2O , THF , reflux, 1.5 – 4 h..

Scheme 28. Possible routes of the reduction of the aromatic nitro group based on the literature.

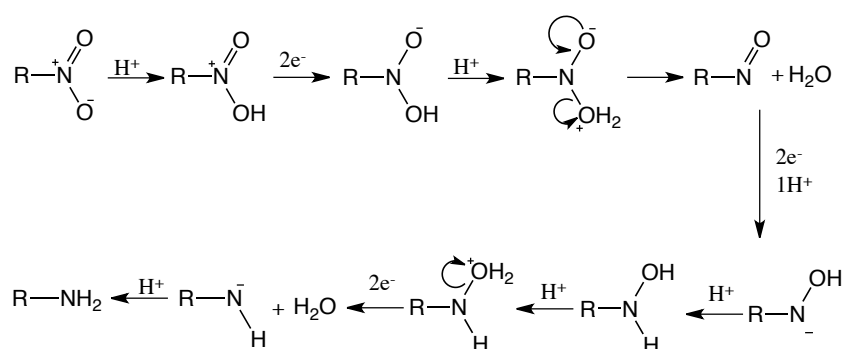


Scheme 29. The reduction of nitro group employing sodium dithionite.



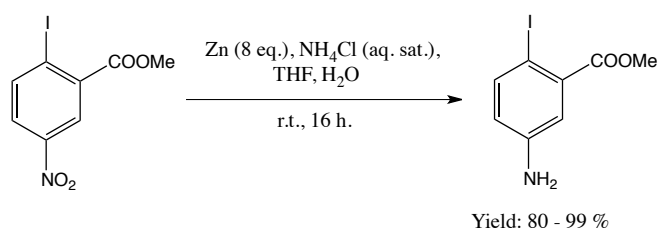
Scheme 30. The reduction of the nitro group employing sodium dithionite.

The reduction of an aromatic nitro group with a metal in an acidic solution is a classical method that is commonly employed in organic synthesis. The accepted mechanism for this type of reduction is a single electron transfer mechanism that can be illustrated as a two-electron transfer, in the following manner:

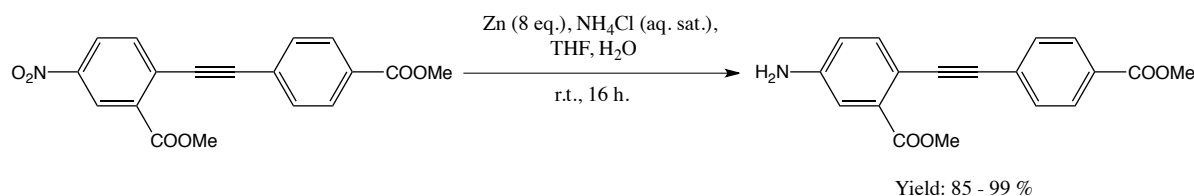


Scheme 31. The accepted mechanism of the reduction of the nitro group using metal in an acidic solution.

The metal functions as an electron donor and the acid as a proton donor. The most commonly used metals are Sn, Fe or Zn. The reduction of the substrates of interest was attempted with FeSO₄ in a water/MeOH system that gave a mixture of products that could not be separated. Reduction with Zn (8 eq.) using a concentrated aqueous solution of NH₄Cl and THF as a solvent furnished the desired products in 45 – 55 % yield. Contrary to this result, most reported procedures using the aqueous Zn/THF/NH₄Cl system had obtained yields over 80 %. Increasing the amount of the metal did not give an improvement in the yields. It was noticed that a yellow slurry was formed during the reaction. It dissolved slightly in THF, EtOAc or CH₂Cl₂. After further investigation of the reaction parameters, it became clear that if the Zn was added at room temperature to a solution of the substrate and the acid, an exothermic reaction resulting in the formation of insoluble compounds ensued, from which the substrate would not undergo further reduction to the amine. If the Zn was added in small portions over a prolonged period of time to a pre-cooled solution of the substrate, the formation of insoluble compounds was minimized, resulting in increased yields. A TLC analysis showed almost complete conversion of the starting material. Yields after flash chromatography were 85 – 95 % (schemes 32 and 33).

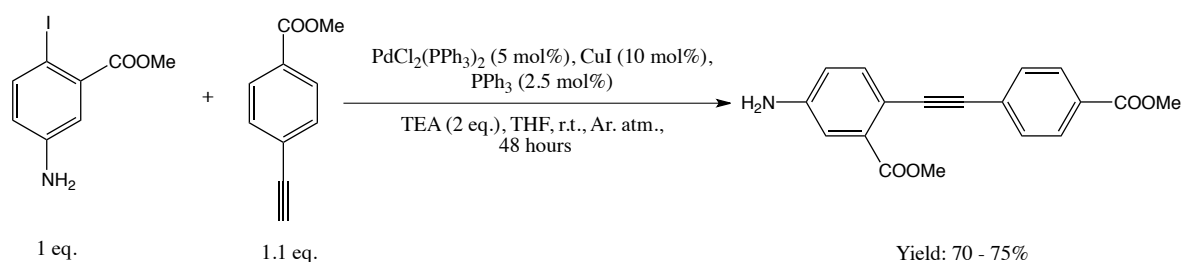


Scheme 32. The reduction of the nitro group employing Zn and sat. aq. NH_4Cl .



Scheme 33. The reduction of the nitro group employing Zn and sat. aq. NH_4Cl .

The Sonogashira coupling between the methyl 4-ethynylbenzoate and the methyl 5-amino-2-iodobenzoate was also tested (scheme 34), furnishing the desired product in 70 – 75 % yield.



Scheme 34. The Sonogashira coupling between the methyl 4-ethynylbenzoate and the methyl 5-amino-2-iodobenzoate.

However, having established an excellent method for the reduction of the nitro group, the attention directed to the hydrolysis of the two methyl ester groups in methyl 5-amino-2-((4-(methoxycarbonyl)phenyl)ethynyl)benzoate. Utilizing 3 eq. of LiOH in THF/water furnished the hydrolyzed product in 90% yield. Our attention was then shifted to the step that proved to be the most challenging – the *cis*-reduction of the triple bond between the two aromatic rings.

10.4.8 Introduction to the selective *cis*-reduction of the triple bond

The reduction of the triple bond could be performed either on compound 1, 2 or 3 (see figure 66). Compound 3, with the carboxylic acid groups, would make purification of a theoretical analogue with a *cis*-double bond by flash chromatography complicated and it was thus not treated as a potential substrate for the *cis*-reduction. Compounds 1 and 2 were perfect substrates for the *cis*-reduction reaction, but the

amino group in the compound 2 could react in an unpredictable manner. Through the rest of the text, the two hydrogens that sit on the double bond will be referred to as H_a and H_b (figure 67).

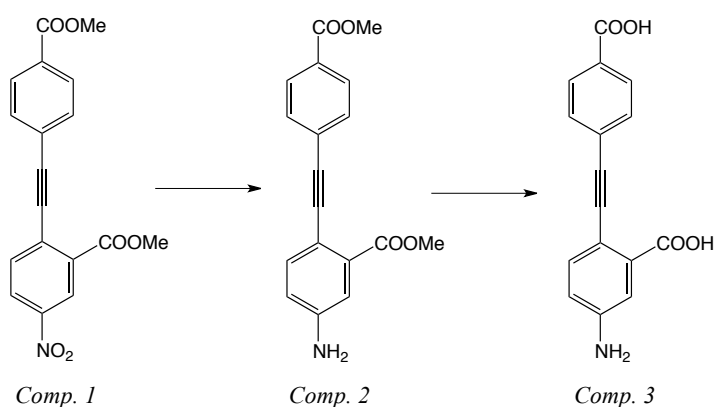


Fig. 66. Substrates of the interest.

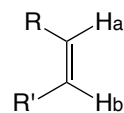
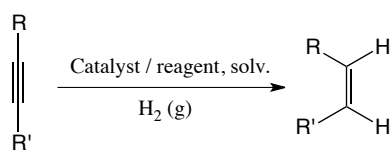
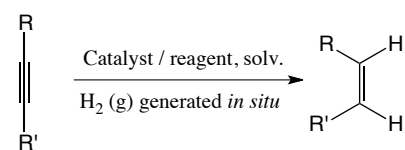


Fig. 67. Nomenclature of the two hydrogens.

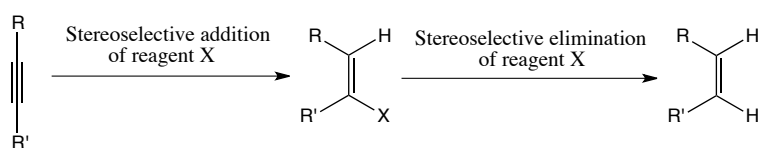
The *cis*-double bond can be synthesized by a variety of methods. It can either be obtained from a triple bond by reduction or it can be synthesized stereoselectively as a *Z*-olefin from suitable substrates. A short overview of the strategies can be found in schemes 35 – 39.



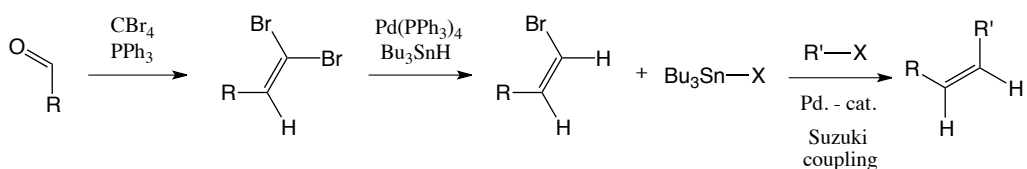
Scheme 35. *Cis*-reduction of the triple bond.



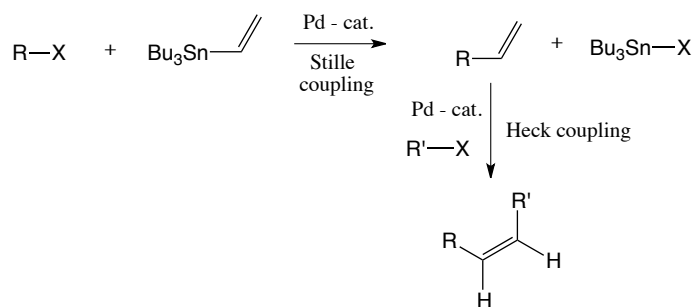
Scheme 36. *Cis*-reduction of the triple bond.



Scheme 37. *Cis*-reduction of the triple bond using an addition/elimination approach.

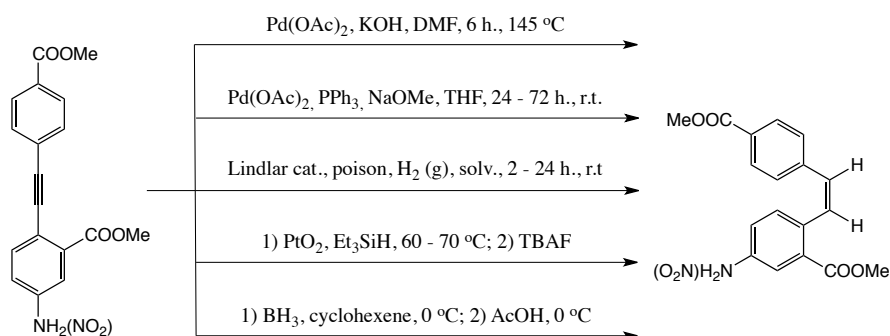


Scheme 38. Stereoselective synthesis of the *cis*-double.



Scheme 39. Two-step synthesis of the *cis*-double bond using Stille and Heck coupling reactions.

The schemes 38 and 39 illustrate a method where the *cis*-double bond is synthesized stereoselectively, avoiding going via a triple bond intermediate which subsequently must be reduced. In the scheme 39, tributyl(vinyl)stannane is being used as a coupling reagent, giving tributyl(bromo or iodo)stannane as a by-product. The same by-product can be found in scheme 38. It is known that it is not trivial to remove this by-product from a reaction mixture. It is highly toxic and since the aim of this project is to prepare the novel pharmaceutical agents, it will not be optimal to have traces of this material in the final product. When it comes to use of the tributyl(vinyl)stannane, extremely rigorous safety routines must be followed. Therefore, the synthesis of the *Z*-olefin linkage by these two protocols was not considered as a safe nor attractive route. The schemes 35 and 36 illustrate the reduction of the triple bond using a catalyst and hydrogen gas. The reaction protocol in the scheme 37 differs from the other procedures since it does not involve a reducing agent – it is more of a “stereoselective addition-elimination reaction”. Based on the literature and the availability of the reagents, the following methods were chosen for testing either on the compound 1 or 2 (scheme 40).

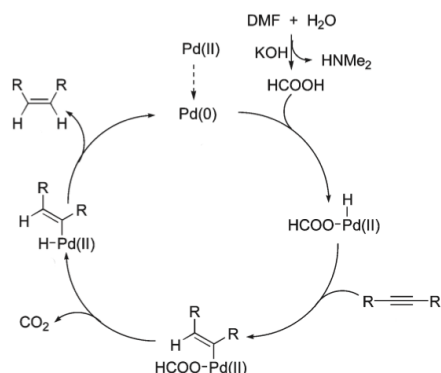


Scheme 40. The *cis*-reduction techniques that were tested either on compound 1 or 2.

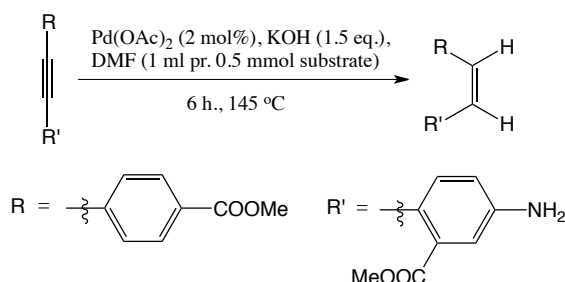
10.4.9 Reduction of the triple bond using the Pd(OAc)₂/KOH/DMF-system

In 2010, an interesting article appeared by Li J. et al.⁵⁵ dealing with the transfer hydrogenation of diarylacetylenes to *cis*-stilbenes. A variety of diaryl compounds with a triple bond were reduced to their *cis*-double bond analogues in excellent yields and with high stereoselectivities. The catalytic

system/reagents used were Pd(OAc)₂/KOH/DMF. According to the article, the proposed mechanism for the transfer hydrogenation reaction is depicted in scheme 41, the optimal stoichiometry and reaction conditions are illustrated in scheme 42.

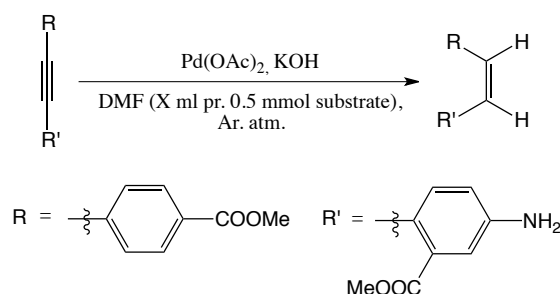


Scheme 41. The mechanism of reduction of the triple bond.



Scheme 42. The optimal stoichiometry and reaction conditions.

Despite of the quite harsh reaction conditions, this protocol was tested on the methyl 5-amino-2-((4-(methoxycarbonyl)phenyl)ethynyl)benzoate. As it can be seen from scheme 41, hydrolysis of the DMF in the presence of KOH generates HCOOH *in situ* in an appropriate concentration that later functions as a hydrogen source. Keeping in mind that the substrate carries two methyl ester groups, it would be right to think that the hydrolysis of the methyl ester groups will be a major side-reaction at such high temperatures. Assuming that the provided mechanism provides an accurate picture of the reaction, the carboxylic acids formed from the hydrolysis of the methyl esters could undergo oxidative addition to the Pd(0) species and further transformations. The first attempts on the *cis*-reduction using an optimized reaction protocol (scheme 42) that according to the authors gave the best yield and stereoselectivity, furnished the desired product in 24 % yield with 70% purity. As it was pointed out in the article, old and undried DMF had been used. TLC-analysis and purification of the crude material by flash chromatography showed that the starting material and the desired product had very similar R_F-values. The desired compound could not be obtained completely pure as per ¹H NMR. Changing the reaction parameters or stoichiometries of reagents/catalyst did not improve the reaction yields (scheme 43, table 8).



Scheme 43. A general procedure for the reduction of the triple bond using the Pd(OAc)₂/KOH/DMF-system.

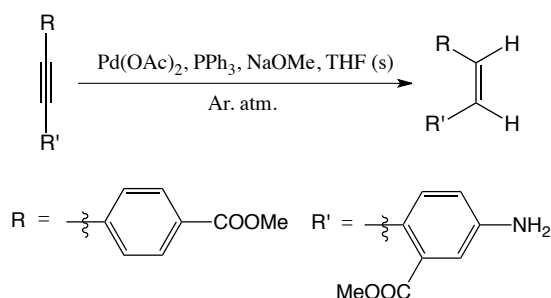
Entry	Pd(OAc) ₂	KOH	DMF (X ml pr.	React. temp.	React. time	Yield:
			0.5 mmol substr.)			
1	2 mol %	1.5 eq.	3	145	6	0 %
2	4 mol %	1.5 eq.	1	145	6	24 %, ca 60 % pure
3	6 mol %	1.5 eq.	1	100	12	26 %, ca 65 % pure

Table 8. The stoichiometries of the reagents and catalyst in the reaction depicted in scheme 43.

Despite the low yields and an extremely tedious work up, some very important facts were unveiled. The most obvious of them is that it is possible to *cis*-reduce the triple bond. An investigation of the data from the ¹H MNR analysis revealed that *J* H_a-H_b in (*Z*)-methyl 5-amino-2-(4-(methoxycarbonyl)-styryl)benzoate had coupling constant of 12.0 – 12.2 Hz (attachment 2, spectra 1 and 2, doublets at 7.073 – 7.033 and 6.540 – 6.500 ppm). The coupling constant was the only proof that the reduction of the triple bond had proceeded according to plan, giving the *cis*-product. Theoretically, the *J* H_a-H_b in a *trans* analogue of (*Z*)-methyl 5-amino-2-(4-(methoxy-carbonyl)styryl)benzoate would have coupling constant of approximately 15 Hz.⁷⁵ Neither the (*Z*)-methyl 5-amino-2-(4-(methoxycarbonyl)styryl)-benzoate nor its *trans* analogue - (*E*)-methyl 5-amino-2-(4-(methoxycarbonyl)- styryl)benzoate were previously reported.

10.4.10 Reduction of the triple bond using the Pd(OAc)₂/PPh₃/NaOMe/THF-system

Another system for the stereoselective reduction of substituted diarylacetylenes was proposed by Wei et al.⁵⁶ The system was tested on the methyl 5-amino-2-((4-(methoxycarbonyl) phenyl)ethynyl)benzoate, unfortunately giving only the starting material, that was recovered in 80 – 90% yield after the flash chromatography (scheme 44, table 9).



Scheme 44. A general procedure for the reduction of the triple bond using the Pd(OAc)₂/PPh₃/NaOMe/THF-system.

Entry	Pd(OAc) ₂	PPh ₃	NaOMe	React. temp. (°C)	React. time (h.)	% recovery of start. mater.
1	5 mol %	5 mol %	5 eq.	r.t.	12	85 – 90

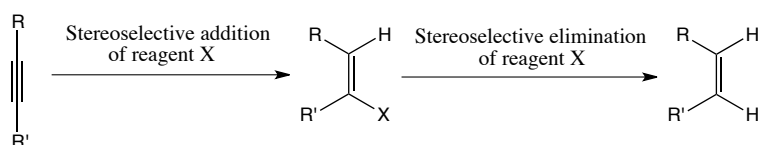
2	10 mol %	10 mol %	10 eq.	r.t.	48	75 – 80
3	10 mol %	10 mol %	10 eq.	45	12	0

Table 9. The stoichiometries of the reagents and catalyst in the reaction depicted in scheme 44.

After observing the difficulty in the reduction of the triple bond using two different protocols, it was decided to employ a distinct approach towards the synthesis of the analogue with the *cis*-double bond.

10.4.11 Synthesis of the *cis*-double bond using an addition/elimination approach

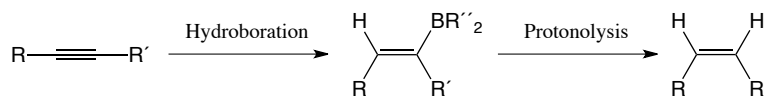
The transformation of a triple bond to a *cis*-double bond can be achieved in two steps using the addition/elimination approach that is illustrated in scheme 45.



Scheme 45. Two-step reduction of the triple bond using an addition/elimination approach.

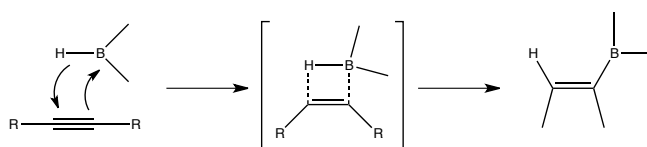
10.4.12 Synthesis of the *cis*-double bond using a hydroboration/protonolysis approach

One of the classical methods for this type of synthesis of a *cis*-double bonds is the hydroboration/protonolysis protocol discovered by Brown et al.⁵⁷ The whole reaction sequence can be viewed in scheme 46.



Scheme 46. A schematic representation of the hydroboration/protonolysis approach.

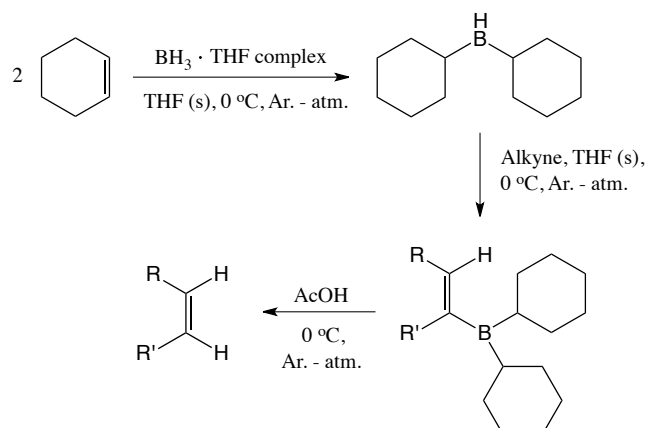
The hydroboration reagent is a disubstituted borane adduct of the type R_2BH . Some of these reagents can be bought from commercial sources; others can easily be prepared *in situ*. The disubstituted borane adduct acts as an electrophile in the hydroboration step since the borane atom has a free *p*-orbital that undergoes a nucleophilic attack by the triple bond. The addition goes through a four-center transition state resulting in a vinylborane adduct (scheme 47).

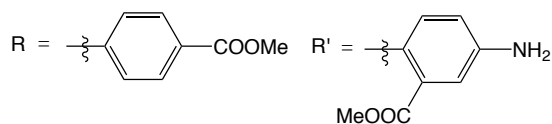


Scheme 47. Mechanism of addition of the disubstituted borane adduct to the triple bond.

The protonolysis of the vinylborane with formic acid proceeds with the retention of the stereochemistry giving a *cis*-olefin. The choice of disubstituted borane adduct to be used in the hydroboration step is governed by the sterical environment of the substrate - heavily hindered alkynes will not react with borane reagents bearing large aromatic groups. The vinylborane adduct itself may react with the hydroborating agent, but this happens only with relatively unhindered alkynes bearing alkane side chains. Both the hydroboration and protonolysis are commonly performed at 0 °C since the vinylborane adducts are not stable and may decompose – as a consequence of this, the vinylborane adducts are not isolated and are used directly in the next step.

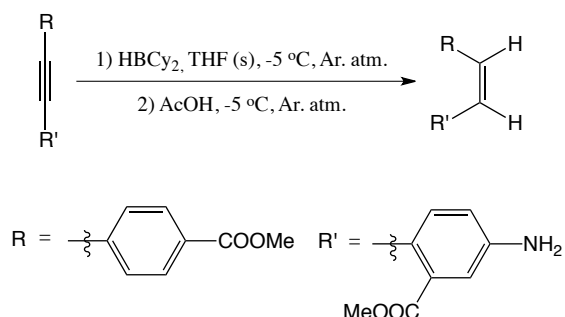
Going back to our substrates of interest, it was decided to adopt the experimental procedures that used dicyclohexylborane as the hydroborating agent.⁵⁸ The hydroboration/protonolysis sequence was tested on the methyl 5-amino-2-((4-(methoxycarbonyl)phenyl)ethynyl)benzoate since it is soluble in THF at low temperatures (down to -20 °C), while the methyl 2-((4-(methoxycarbonyl)phenyl)ethynyl)-5-nitrobenzoate is not. During the hydroboration step, the disubstituted borane reagent reacts in an *anti*-Markovnikov fashion, adding to the most hindered carbon of the two that constitute the triple bond. For our purposes, addition to either of the carbons in the triple bond would suffice. During the protonolysis step with formic acid, hydrolysis of the methyl ester groups was a possible side reaction, but since the step is commonly done at 0 °C, the hydrolysis would be very slow or should not take place at all. The overall reaction sequence can be viewed in scheme 48.





Scheme 48. A general procedure for the synthesis of the *cis*-double bond using the hydroboration/protonolysis approach.

As can be seen from the scheme 48, the dicyclohexylborane was generated *in situ* from the cyclohexene and the $\text{BH}_3 \cdot \text{THF}$ complex. Unfortunately, a ^1H NMR analysis of the crude mixture showed no formation of the desired product, the starting material was always recovered after flash chromatography in 90 – 95% yield. Attempts with larger excesses of the dicyclohexylborane, prolonged reaction times etc. (scheme 49, table 10) did not give any conversion of the starting material. An interesting observation is that in all cases, the white solid (dicyclohexylborane) disappeared after the addition of the substrate – maybe due to formation of the vinylborane adduct, but since there was no product formed, this interpretation goes unsubstantiated. No hydrolysis of the methyl ester groups was observed in any of the cases.



Scheme 49. A general procedure for the synthesis of the *cis*-double bond using the hydroboration/protonolysis approach.

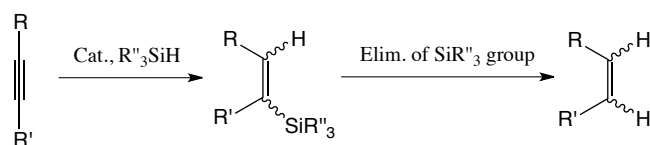
Entry	HBCy ₂	AcOH	React. time (hydrobor. + protonolysis, h)	% conversion to alkenyl adduct
1	2 eq.	12 eq.	1 + 0.5	0
2	4 eq.	24 eq.	2 + 1	0
3	10 eq.	60 eq.	2 + 2	0
4	15 eq.	80 eq.	2 + 3	0
5	20 eq.	120 eq.	3 + 2	0

Table 10. The stoichiometries of the reagents and catalyst in the reaction depicted in scheme 49.

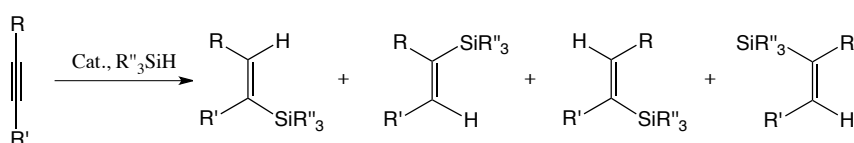
10.4.13 Synthesis of the *cis*-double bond using hydrosilylation/protodesilylation approach

The next approach that was attempted to effect the *cis*-reduction of the triple bond was based on the research done by Hamze A. et al.⁵⁹ The main idea was exactly the same as with the hydroboration/

protonolysis approach – an addition of a H–Si bond to the *sp*-carbons in the triple bond in a *cis*-selective fashion, followed by the removal of the silane adduct, leaving behind a *cis*-double bond (scheme 50). This approach was used in the synthesis of analogues of combretastatin.^{59a} An addition of a H–Si bond to the carbon-carbon triple bond can give 4 different products (scheme 51).

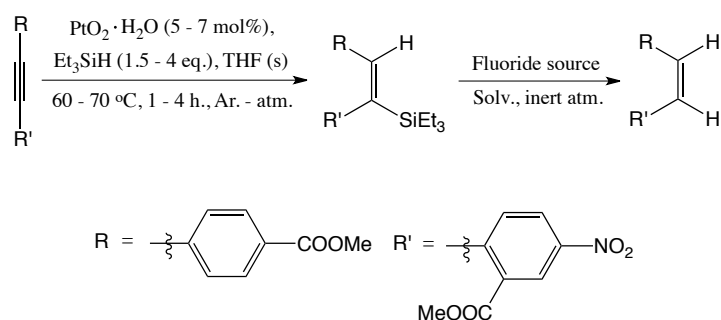


Scheme 50. The main concept behind the hydrosilylation/protodesilylation approach.



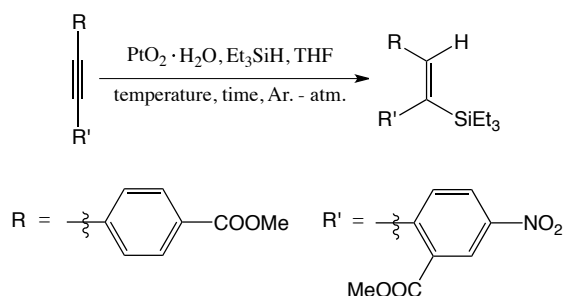
Scheme 51. Possible product distribution in the hydrosilylation reaction.

Only (*E*)-alkenylsilanes will lead to the formation of the *cis*-alkene after elimination of the silyl group. In their research Hamze et al. employed platinum dioxide as a catalyst in the hydrosilylation of substituted diarylacetylenes, giving only (*E*)-adducts. Which of the two (*E*)-alkenylsilanes is being formed is not crucial for this project since both give *cis*-alkenes after the elimination of the silyl group. The carbon-silicon bonds are readily cleaved by TBAF or other sources of fluoride ions.^{59a, 60} The alkenylsilanes are isolable compounds by flash chromatography and have been characterized by NMR.^{59c} NOESY techniques have been used to prove the (*E*)-geometry of the silicon adduct. For our purposes, it meant that the alkenylsilanes could be isolated, purified and characterized before using them in the next step. Cleavage of the carbon-silicon bond by fluoride ions is a known reaction – the stereoselectivity of the cleavage, isomerisation and side-reactions were the major concerns. Keeping all that in mind, the reaction protocol shown in scheme 52 was tested on the methyl 2-((4-(methoxycarbonyl)phenyl)ethynyl)-5-nitrobenzoate (scheme 52).



Scheme 52. A general hydrosilylation/protodesilylation reaction sequence.

In the original procedure that was adopted for the hydrosilylation/protodesilylation of our substrate of interest, the hydrosilylation was done in neat triethylsilane, the silane functioning both as the solvent and a reagent. When the original procedure was tested on the methyl 2-((4-(methoxycarbonyl)-phenyl)-ethynyl)-5-nitrobenzoate, spontaneous evolution of gas and a small explosion accompanied by a small fire ensued upon addition of the triethylsilane. Dissolving the substrate and the catalyst in THF prior to the addition of the triethylsilane did not result in a similarly uncontrolled reaction. The hydrosilylation reaction was tested with different catalyst loadings, amounts of triethylsilane, reaction times and temperatures (scheme 53). None of the reaction outcomes were reproducible – the same conditions gave very different outcomes in terms of yields of the desired product – anything from no conversion to 86% yield was obtained. It is unclear why this happened. In cases where the reaction worked fine, one could see the bubbles coming from the reaction mixture, most probably hydrogen gas.⁶¹ Evolution of the hydrogen gas may also explain the spontaneous combustion when the reaction was performed neat. Platinum metal reacts very violently with hydrogen gas and since the alkyne substrate is not soluble in triethylsilane, the reaction mixture was not homogeneous – the heat and energy did not have any place to escape. In the references there was no information about this –perhaps indicating that similar situations did not occur in their cases.^{59, 60} Most likely, the alkyne substrates tested were in liquid form, thereby making the process of the heat transport from the reaction mixture to its surroundings more efficient. When the reaction did not work (conversion < 5 %), no bubbles were observed. On the positive side, the starting material and the alkenylsilanes adduct could be easily separated by flash chromatography, making the process of screening more economical and efficient. Despite the failure of the hydrosilylation reaction, the cleavage of the carbon-silicon bond was examined. TBAF was used as a source of fluoride ions together with copper iodide acting as a promoter of the protodesilylation reaction and buffering the activity of the TBAF.⁶² A summary of the hydrosilylation/protodesilylation experiments is presented in scheme 53, table 11, scheme 54 and table 12.

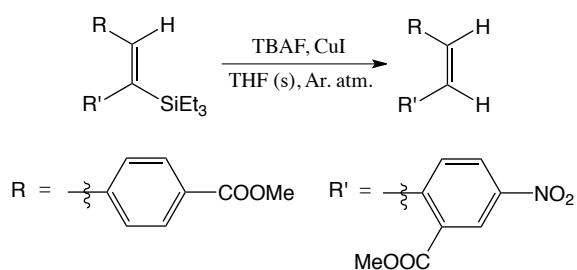


Scheme 53. A general protocol for the hydrosilylation reaction.

Entry	PtO ₂ hydrate (mol %)	Et ₃ SiH (eq.)	React. temp. (°C)	React. time (h)	% conv. to alkenylsil. adduct
1	7	1.5	60	2	80

2	7	1.5	60	2	< 5
3	20	1.5	60	2	< 5
4	10	10	60	2	0
5	7	5	60	2	< 5
6	15	1.5	70	2	< 5
7	3	3	56	2.5	50
8	6.5	3	55	3.5	< 5
9	6.5	3	55	3.5	40 %
10	5	3	58	3	64 %
11	5	3	58	3	< 5
12	5	1.6	56	3	< 5

Table 11. The stoichiometries of the reagents and catalyst in the reaction depicted in scheme 53.



Scheme 54. A general protocol for the protodesilylation reaction.

Entry	TBAF (eq.)	CuI (eq.)	React. t. (°C)	React. time (h)	Cis-trans ratio
1	3.2	0	0	6	<i>trans</i> only
2	2	0.2	r.t.	16	<i>Cis</i> (64 %) + <i>trans</i> (36 %)
3	1.1	1.5	- 30	16	<i>Cis</i> (40 %) + <i>trans</i> (60 %)
4	3	0	- 5	2	<i>trans</i>

Table 12. The stoichiometries of the reagents and catalyst in the reaction depicted in scheme 54.

The TBAF was dried prior to use by dissolving it in dry toluene, concentrating *in vacuo*, and then drying for 3 days under high vacuum. As it can be seen from scheme 54, table 12, the protodesilylation of the alkenylsilane adduct with TBAF resulted in a mixture of the *cis/trans* products or only *trans* products. The yields in all cases were low – around 20 % - suggesting that a side-reaction might be involved. In cases where only TBAF was used, only the *trans* product was obtained – TBAF most probably cleaved the carbon-silicon bond in a stereoselective fashion, giving only the *cis* alkene in the first place and then isomerized it to the *trans* alkene. In cases where the CuI was used as a “buffer”, the *cis* isomer was also present. Low temperatures did not make a particular impact on a rate of the isomerization etc. The *trans* and *cis* analogues could not be separated by flash chromatography. The $J_{\text{H}_a\text{-H}_b}$ in the *cis* analogue was 12.2 Hz (attachment 2, spectra 5, 7, doublets at 7.192 – 7.138 and 6.858 – 6.817 ppm; attachment 2,

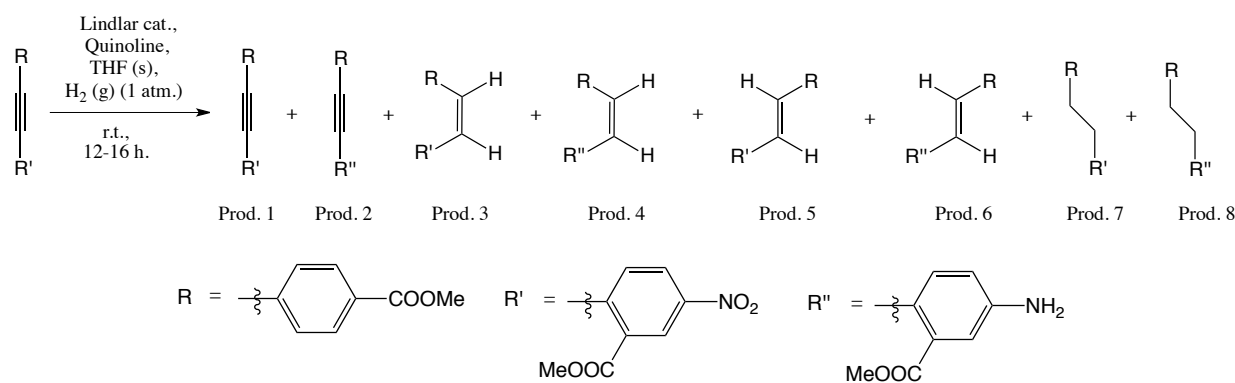
spectra 3, 4, doublets at 7.198 – 7.136 and 6.866 – 6.805 ppm), the $J_{H_a-H_b}$ in the *trans* analogue was 16.2 Hz. (attachment 2, spectra 8 – 11, doublets at 8.204 – 8.149 ppm and 7.187 – 7.133 ppm). In the reactions where the *trans* adduct was the only product, the product was a yellow solid. When mixtures of both the *cis* and *trans*-isomers were obtained, the purified product consisted of both the yellow solid and an oil, suggesting that the *cis*-isomer is a yellow oil.

10.4.14 Synthesis of the *cis*-double bond by hydrogenation using Lindlar catalyst system

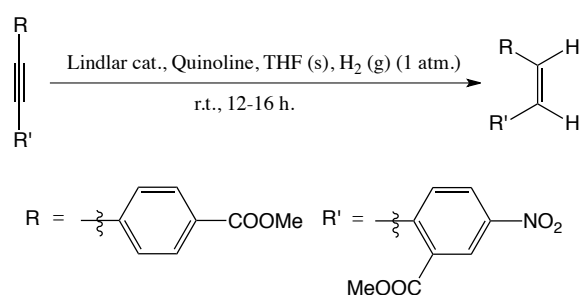
The Lindlar catalyst is a catalyst system that is widely used for partial hydrogenation of acetylenic systems to alkenes.⁶³ Alkyne reduction is a stereoselective process, occurring via *syn* addition to give *cis* alkenes. The Lindlar catalyst is a heterogeneous catalyst system that consists of 5 %wt palladium deployed on a porous calcium carbonate surface with a deactivating agent – lead or quinoline. The selectivity of the partial hydrogenation of alkynes can be divided into three stages/parts: hydrogenation stops at the alkene stage, no other functional groups undergo reactions and reactant/product(s) of the reaction do not undergo oligomerization reactions.

The reduction of the triple bond with the Lindlar system could be performed on the two substrates of the interest – either on the methyl 2-((4-(methoxycarbonyl)phenyl)ethynyl)-5-nitrobenzoate or the methyl 5-amino-2-((4-(methoxycarbonyl)phenyl)ethynyl)benzoate. The amino group could potentially lead to over-reduction of the alkyne moiety to the alkane. Protection of the amino group is a common strategy to overcome this issue, but the introduction of the two additional steps would be of a little practical value. The nitro group itself could be reduced to the amino group under the reaction conditions that are typical for hydrogenations with the Lindlar catalyst.⁶⁴

Control over the Lindlar catalyst and its activity can be achieved through the use of poison – lead or quinoline. The Lindlar catalyst that was available already contained lead acetate, so quinoline was employed as an additional deactivator of the catalyst. Quinoline acts by blocking the potential adsorption sites for dihydrogen and the alkyne.⁶⁵ Hydrogen-bonding solvents can also be employed as deactivators of the catalyst,⁶⁶ but poor solubility of the methyl 2-((4-(methoxycarbonyl)phenyl)ethynyl)-5-nitrobenzoate in MeOH or EtOH made it impossible to use this technique. The methyl 2-((4-(methoxycarbonyl)phenyl)ethynyl)-5-nitrobenzoate became the substrate that was tested in partial hydrogenation reactions employing the Lindlar catalyst system. The idea was to start with a big percentage of the quinoline and observe what would happen when the ratio between the catalyst/poison was changed. The hydrogen gas pressure was 1 atmosphere, because higher pressures could eventually lead to unwanted side reactions such as over-reduction. Theoretically, the reduction could give 8 different products (scheme 55). An overview of the experiments and product distributions can be viewed in scheme 56, table 13.



Scheme 55. Probable products after the hydrogenation reaction with Lindlar catalyst.



Scheme 56. A general reaction protocol for the reduction of the triple bond employing Lindlar catalyst.

Entry	Lindlar cat. (mol%)	Quinoline (mol%)	Prod. distrib. (nomenclature as in scheme 58):
1	5	30	Prod. 1 + prod. 3
2	5	15	Prod. 2 + prod. 4
3	5	15	Prod. 4
4	5	10	Prod. 2
5	5	5	Prod. 2

Table 13. The stoichiometries of the reagents and catalyst in the reaction depicted in scheme 56.

In all cases, hydrogen gas was introduced into the reaction flask from a balloon via a hypodermic needle that was stuck through the septum. The reaction flask was flooded with hydrogen. Hydrogen gas was also bubbled through the solvent, after that the gas inlet was removed and the flask was sealed with plastic tape. Using 30 mol% of quinoline (entry 1, table 13), furnished a mixture of the starting material and the product 3 (scheme 58). They could not be separated by flash chromatography. The $J_{H_a-H_b}$ in product 3 had value of 12.4 Hz (attachment 2, spectra 12, 13, doublets at 7.200 – 7.139 and 6.870 – 6.808 ppm). Upon lowering the amount of quinoline to 15 mol% (entries 2 and 3, table 13) reduction of both the triple bond and the nitro group was observed (attachment 2, spectra 14, 15). Upon repetition of the reaction, only product 4, with a *cis*-double bond and an amino, was formed and isolated in 96 %

yield. The $J_{H_a-H_b}$ in product 4 had value of 12.0 Hz. Further lowering of the amount of quinoline (entries 4 and 5, table 13) resulted only in the reduction of the nitro group, giving only product 2.

Summarizing the reactions in table 13, several things can be noted. The formation of the *trans* product was not observed in any cases, meaning that the reduction proceeded in a highly *cis*-selective fashion and that no isomerization occurred. Upon lowering the quantity of quinoline, reduction of the nitro group was the only reaction – reduction of the triple bond did not take place at all (entries 4, 5, table 13). Assuming that preassociation of the alkyne bond to the metal surface is one of the first steps in the reduction of the triple bond, it is unclear why only reduction of the nitro group occurred. As it was pointed out in the literature,⁶⁶ the amino group is usually protected to avoid over-reduction to an alkane, while in this case not even reduction to an alkene was observed. Formation of the amino group somehow lowered the activity of the catalyst, making it less efficient towards reduction of the triple bond.

Chapter 11

Conclusion

Compounds AB74, 75, 76, 77, 78 and 81 have been docked against the DBD of the c-Myb, synthesized and tested in the EMSA analysis. The biological analysis revealed severe limitations of the AutoDock 4.2 software when it came to the prediction of the inhibition constants.

Compound NSC675460 has been successfully modified using the bioisosteric approach, resulting in three new compounds with better *in silico* activities than NSC675460.

The synthesis of the compound NSC675460 could not be finished due to the short time that was left at the end of the project. The hydrolysis of the ester groups and the azo-coupling reaction with the 8-hydroxyquinoline-5-sulfonic acid remains to be investigated.

Chapter 12

Experimental

General

Chemicals were used as delivered from Sigma-Aldrich or TCI Europe. Argon (4.0) was used as delivered from AGA.

Thin layer chromatography was done on 60 F₂₅₄ silica coated aluminium plates from Merck. Flash chromatography was done on silica gel Merck (Silikagel 60, 0.40-0.63 mm, 480-540 m²/g, pH 6.5 – 7.5) with an Isco Inc. CombiFlash Companion with PeakTrak software (v. 1.4.10). Ethyl acetate and hexane used as eluents were of technical grade.

NMR spectra were recorded on a Bruker Avance DPX200, DPX300 and DPX400 instrument at 200, 300 or 400 MHz for ¹H-NMR and 50, 75 and 100 MHz for ¹³C-NMR at 25 °C. Chemical shift (δ) is given in ppm relative to the solvent CDCl₃ (7.24 ppm for ¹H-NMR, 77.0 ppm for ¹³C-NMR), DMSO (2.49 ppm for ¹H-NMR, 39.50 ppm for ¹³C-NMR). Coupling constants are given in Hz. All ¹³C-NMR spectra are decoupled. Peak assignment in ¹H- and ¹³C-NMR spectra is based on information from DEPT135, DEPT 90, HMQC and COSY experiments as needed. Peaks due to solvents such as dichloromethane, ethyl acetate, water, hexane and THF are not included in the interpretation of NMR spectra.

MS spectra are recorded on a VG Prospec sector instrument from Fissions Instruments at 70 eV (EI). HR-MS is recorded using perfluorokerosene (PFK) as reference. For electrospray (ESI) a Micromass Q-Tof-2 mass spectrometer was used.

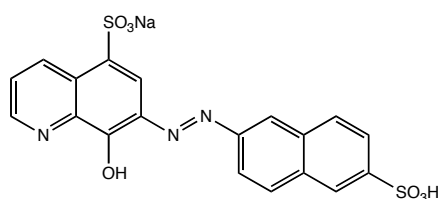
General procedure for the synthesis of compounds AB 74, 75, 76, 77, 78 and 81

To a round-bottomed flask, amine (0.483 g, 2 mmol, 1 eq.), Na₂CO₃ (0.101 g, 1 mmol, 0.5 eq.) and H₂O (12 mL) were added. The mixture was stirred for 20 min. at r.t. and then cooled down to 0°C. Then conc. HCl (0.34 mL) in H₂O (2 mL) was added and the solution was stirred for additional 20 min at 0°C. Then a solution of NaNO₂ (0.138g, 2 mmol, 1 eq.) in H₂O (0.322 ml) was added slowly at 0°C over 10 min. The reaction mixture was stirred at 0°C for 1 hour and then added to a solution of 8-hydroxyquinoline-5-sulfonic acid monohydrate (0.486 g, 2.0 mmol, 1 eq.) in 1M NaOH solution (6

mL). The reaction mixture turned purple and was stirred for 3 hours at 0°C. Cold conc. HCl (25 droplets) and NaCl (0.75 g) were added and the product precipitated after 30 min. The product was vacuum filtered and collected. The crude product was dried overnight at r.t. Purification was done by dissolving crude product in minimum amount of 1M NaOH, filtrating it and then precipitating product by careful addition of 12M HCl. The precipitated purified product was collected by vacuum filtration. The product was purified two times by the same method.

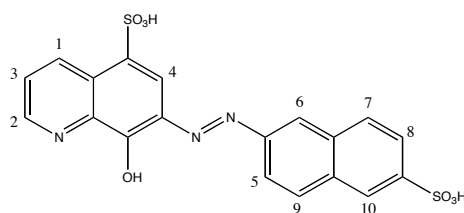
Purity of the compounds was checked by HPLC using a Spectra SYSTEM P2000 pump with a Spectra SYSTEM UV3000 UV detector with a reverse phase column – Kromasil 100-5-C18, 250 x 10 mm. Eluent system was acetonitrile / ammonium acetate buffer (pH=10), flow = 2,5 ml/min, acetonitrile 0→100% over 20 min.

Synthesis of sodium (*E*)-8-hydroxy-7-((6-sulfonaphthalen-2-yl)diazenyl)quinoline-5-sulfonate

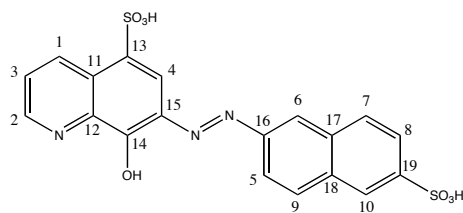


The compound was prepared and purified according to the general procedure.

Yield: 0.2073 g (22%)



¹H NMR (300 MHz, DMSO-*d*₆ as solvent, in ppm): δ 9.554 - 9.525 (1H, d, J = 8.7 Hz, **H1**); 9.083 - 9.070 (1H, d, J = 3.9 Hz, **H2**); 8.677 (1H, s, **H10**); 8.481 (1H, s, **H4**); 8.310 - 8.275 (1H, dd, J = 1.8 Hz, J = 8.7 Hz, **H8**); 8.244 (1H, s, **H6**); 8.152 - 8.107 (2H, dd, J = 4.8 Hz, J = 8.7 Hz, **H7**, **H9**); 8.062 - 8.017 (1H, dd, J = 4.8 Hz, J = 8.7 Hz, **H3**); 7.847 - 7.814 (1H, dd, J = 1.5 Hz, J = 8.7, **H5**)



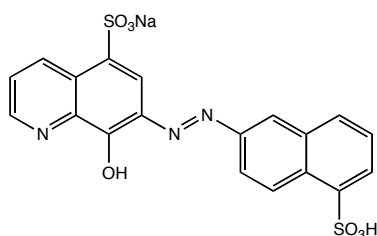
^{13}C NMR (75 Mhz, DMSO- d_6 as solvent, in ppm): δ 152.01 (Ar-C); 149.36 (Ar-C); 147.34 (Ar-C); 145.88 (C2); 142.71 (C1); 135.46 (Ar-C); 135.22 (Ar-C); 134.98 (Ar-C); 133.74 (Ar-C); 132.97 (Ar-C); 130.06 (C7); 129.18 (C9); 127.85 (Ar-C); 127.43 (C10); 125.10 (Ar-C); 124.30 (C6); 124.14 (C3); 117.55 (C8); 116.87 (C4).

MS (ESI, TOF MS ES $^-$) m/z (relative intensity): 480.1 (100%)

HR-MS (ESI, TOF MS ES $^-$): Calculated value for $\text{C}_{19}\text{H}_{11}\text{N}_3\text{NaO}_7\text{S}_2^-$: 479.9936. Found: 479.9925 (-2.32 ppm)

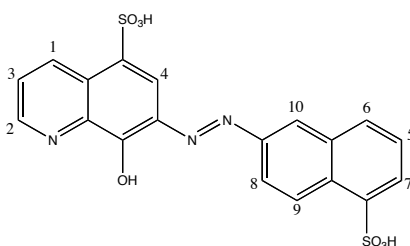
HPLC: 7.8 min; 93% pure

Synthesis of sodium (*E*)-8-hydroxy-7-((5-sulfonaphthalen-2-yl)diazenyl)quinoline-5-sulfonate

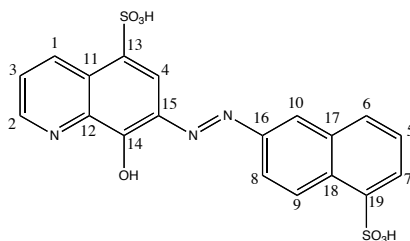


The compound was prepared and purified according to the general procedure.

Yield: 0.4958 g (34%)



¹H NMR (300 MHz, DMSO-d₆ as solvent, in ppm): δ 9.610 - 9.577 (1H, dd, J = 1.2 Hz, J = 8.7 Hz, **H1**); 9.125 - 9.104 (1H, dd, J = 1.2 Hz, J = 4.8 Hz, **H2**); 8.996 - 8.965 (1H, d, J = 9.3 Hz, **H9**); 8.707 - 8.701 (1H, d, J = 1.8 Hz, **H10**); 8.503 (1H, s, **H4**); 8.314 - 8.276 (1H, dd, J = 2.1 Hz, J = 9.3 Hz, **H8**); 8.162 - 8.049 (3H, m, **H7, H6, H3**); 7.597 - 7.546 (1H, t, J = 7.8 Hz, **H5**).



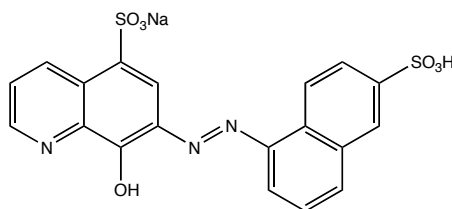
¹³C NMR (75 Mhz, DMSO-d₆ as solvent, in ppm): δ 152.06 (**Ar-C**); 148.46 (**Ar-C**); 145.94 (**C2**); 144.36 (**Ar-C**); 142.70 (**C1**); 135.46 (**Ar-C**); 135.23 (**Ar-C**); 135.13 (**Ar-C**); 133.90 (**Ar-C**); 131.05 (**Ar-C**); 130.51 (**Ar-C**); 129.08 (**Ar-C**); 127.85 (**Ar-C**); 127.58 (**Ar-C**); 126.39 (**Ar-C**); 125.90 (**Ar-C**); 124.11 (**C3**); 116.93 (**Ar-C**); 116.71 (**C4**).

MS (ESI, TOF MS ES⁻) *m/z* (relative intensity): 480.1 (100%)

HR-MS (ESI, TOF MS ES⁻): Calculated value for C₁₉H₁₁N₃NaO₇S₂⁻: 479.9936. Found: 479.9925 (-2.31 ppm)

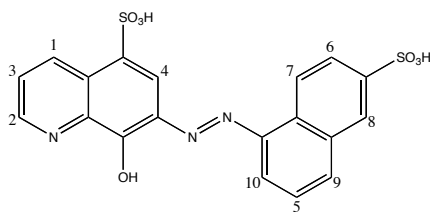
HPLC: 7.3 min; 91% pure

Synthesis of sodium (*E*)-8-hydroxy-7-((6-sulfonaphthalen-1-yl)diazenyl)quinoline-5-sulfonate

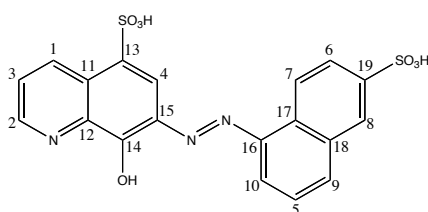


The compound was prepared and purified according to the general procedure.

Yield: 0.2297 g (24%)



¹H NMR (300 MHz, DMSO-d₆ as solvent, in ppm): δ 9.712 - 9.684 (1H, d, J = 8.4 Hz, **H1**); 9.145 - 9.130 (1H, d, J = 4.5 Hz, **H2**); 8.806 - 8.776 (1H, d, J = Hz, J = 9 Hz, **H6**); 8.606 (1H, s, **H4**); 8.362 (1H, s, **H8**); 8.224 - 8.043 (3H, m, J = 7.9 Hz, **H9**, **H10**, **H3**); 8.072 - 8.043 (1H, d, J = 8.7 Hz, **H7**); 7.703 - 7.650 (1H, t, J = 8.15 Hz, **H5**).



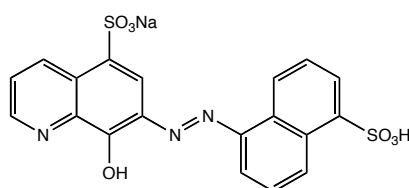
¹³C NMR (75 MHz, DMSO-d₆ as solvent, in ppm): δ 151.67 (**Ar-C**); 145.95 (**Ar-C**); 145.78 (**Ar-C**); 145.55 (**C2**); 144.07 (**C1**); 136.27 (**Ar-C**); 135.34 (**Ar-C**); 133.89 (**Ar-C**); 133.32 (**Ar-C**); 133.06 (**C10**); 130.15 (**Ar-C**); 128.01 (**Ar-C**); 126.71 (**C5**); 125.62 (**C6**); 125.02 (**C8**); 124.57 (**C9**); 122.26 (**C7**); 117.24 (**C4**); 114.18 (**C3**).

MS (ESI, TOF MS ES⁻) *m/z* (relative intensity): 480.1 (100%)

HR-MS (ESI, TOF MS ES⁻): Calculated value for C₁₉H₁₁N₃NaO₇S₂⁻: 479.9936. Found: 479.9925 (-2.29 ppm)

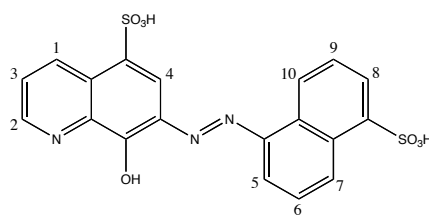
HPLC: 7.6 min; 95% pure

Synthesis of sodium (*E*)-8-hydroxy-7-((5-sulfonaphthalen-1-yl)diazenyl)quinoline-5-sulfonate

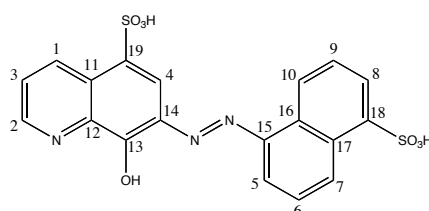


The compound was prepared and purified according to the general procedure.

Yield: 0.3421 g (32%)



¹H NMR (300 MHz, DMSO-d₆ as solvent, in ppm): δ 9.674 - 9.641 (1H, dd, J = 1.5 Hz, J = 8.7 Hz, **H1**); 9.136 - 9.103 (2H, m, **H2, H8**); 8.898 - 8.870 (1H, d, J = 8.4 Hz, **H5**); 8.49 (1H, s, **H4**); 8.229 - 8.206 (1H, d, J = 6.9 Hz, **H7**); 8.166 - 8.120 (2H, m, J = 7.2 Hz, **H10, H3**); 7.774 - 7.712 (2H, dt, J = 2.4 Hz, J = 7.2 Hz, **H6, H9**).



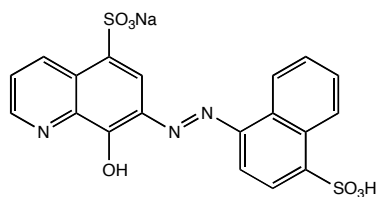
¹³C NMR (75 Mhz, DMSO-d₆ as solvent, in ppm): δ 151.31 (**Ar-C**); 146.28 (**Ar-C**); 145.77 (**C2**); 144.38 (**Ar-C**); 143.36 (**C1**), 136.17 (**Ar-C**); 135.58 (**Ar-C**); 134.43 (**Ar-C**); 132.06 (**C8**); 131.10 (**Ar-C**); 129.94 (**Ar-C**); 127.79 (**Ar-C**); 126.30 (**C6**); 125.69 (**C9**); 125.18 (**C3**); 124.17 (**C10**); 123.72 (**C5**); 116.49 (**C4**); 112.78 (**C7**)

MS (ESI, TOF MS ES⁻) *m/z* (relative intensity): 480.1 (100%)

HR-MS (ESI, TOF MS ES⁻): Calculated value for C₁₉H₁₁N₃NaO₇S₂⁻: 479.9936. Found: 479.9925 (-2.33 ppm)

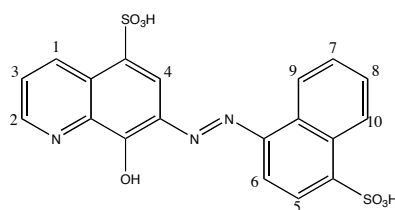
HPLC: 6.9 min; 90% pure

Synthesis of sodium (*E*)-8-hydroxy-7-((4-sulfonaphthalen-1-yl)diazenyl)quinoline-5-sulfonate

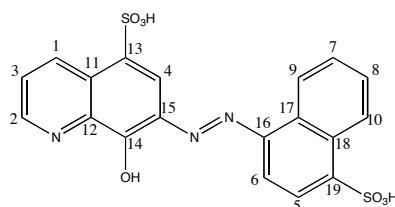


The compound was prepared and purified according to the general procedure.

Yield: 0.395 g (33%)



¹H NMR (300 MHz, DMSO-d₆ as solvent, in ppm): δ 9.653 - 9.626 (1H, d, J = 8.1 Hz, **H1**); 9.141 - 9.128 (1H, d, J = 3.9 Hz, **H2**); 9.020 – 8.992 (1H, d, J = 8.4 Hz, **H9**); 8.873 - 8.845 (1H, d, J = 8.4 Hz, **H10**); 8.55 (1H, s, **H4**); 8.146 - 8.123 (3H, m, **H6**, **H5**, **H3**); 7.815 - 7.767 (1H, t, J = 7.2 Hz, **H7**); 7.720 - 7.672 (1H, t, J = 7.5 Hz, **H8**).



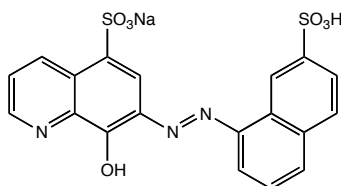
¹³C NMR (75 Mhz, DMSO-d₆ as solvent, in ppm): δ 151.95 (**Ar-C**); 148.00 (**Ar-C**); 147.45 (**Ar-C**); 146.70 (**C2**); 143.94 (**C1**); 136.96 (**Ar-C**); 136.39 (**Ar-C**); 135.65 (**Ar-C**); 131.63 (**Ar-C**); 130.78 (**Ar-C**); 128.96 (**Ar-C**); 128.75 (**C9**); 128.00 (**C8**); 127.30 (**C7**); 125.34 (**C5**); 125.09 (**C6**); 123.00 (**C10**); 117.38 (**C4**); 112.47 (**C3**)

MS (ESI, TOF MS ES⁻) *m/z* (relative intensity): 480.1 (100%)

HR-MS (ESI, TOF MS ES⁻): Calculated value for C₁₉H₁₁N₃NaO₇S₂⁻: 479.9936. Found: 479.9925 (-2.32 ppm)

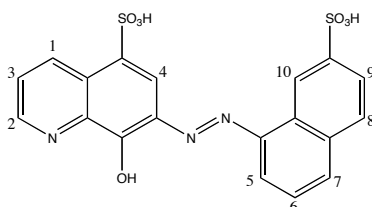
HPLC: 7.4 min; 93% pure

Synthesis of sodium (*E*)-8-hydroxy-7-((7-sulfonaphthalen-1-yl)diazenyl)quinoline-5-sulfonate

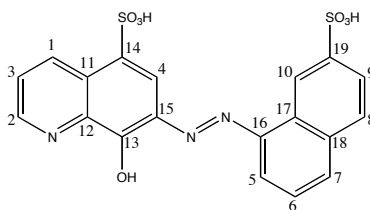


The compound was prepared and purified according to the general procedure.

Yield: 0.1973 g (21.5%)



¹H NMR (300 MHz, DMSO-d₆ as solvent, in ppm): δ 9.750 - 9.724 (1H, d, J = 7.8 Hz, **H1**); 9.191 - 9.176 (1H, d, J = 4.5 Hz, **H2**); 9.056 (1H, s, **H10**); 8.528 (1H, s, **H4**); 8.268 - 8.216 (2H, m, **H5**, **H7**); 8.182 - 8.155 (1H, d, J = 8.1 Hz, **H3**); 8.081 - 8.053 (1H, d, J = 8.4 Hz, **H8**); 7.917 - 7.884 (1H, dd, J = 1.5 Hz, J = 8.4 Hz, **H9**); 7.747 - 7.695 (1H, t, J = 7.8 Hz, **H6**).



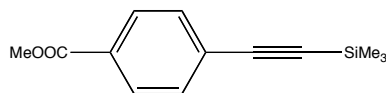
¹³C NMR (75 Mhz, DMSO-d₆ as solvent, in ppm): δ 152.46 (Ar-C); 146.73 (Ar-C); 146.17 (Ar-C); 145.07 (C1); 144.71 (C2); 136.03 (Ar-C); 135.49 (Ar-C); 133.78 (Ar-C); 133.63 (Ar-C); 131.82 (C3); 128.96 (Ar-C); 128.34 (Ar-C); 128.25 (C8); 126.62 (C6); 124.85 (C7, C5); 124.74 (C9); 118.67 (C10); 118.03 (C4); 114.14 (C7, C5).

MS (ESI, TOF MS ES⁻) *m/z* (relative intensity): 480.1 (100%)

HR-MS (ESI, TOF MS ES⁻): Calculated value for C₁₉H₁₁N₃NaO₇S₂⁻: 479.9936. Found: 479.9925 (-2.34 ppm)

HPLC: 7.5min; 92% pure.

Synthesis of methyl 4-((trimethylsilyl)ethynyl)benzoate



Procedure 1

To a round bottomed flask, methyl 4-iodobenzoate (0.524 g, 2 mmol, 1 eq.), PdCl₂(PPh₃)₂ (0.0702 g, 0.1 mmol, 5 mol%), CuI (0.0381 g, 0.2 mmol, 10 mol%), PPh₃ (0.0131 g, 0.05 mmol, 2.5 mol%) were added, the flask was equipped with septum and flushed with argon. Then dry and degassed THF (5 ml) was added with a syringe. Then, a deoxygenated solution of ethynyltrimethylsilane (0.424 ml, 3 mmol, 1.5 eq.) in TEA (0.418 ml, 3 mmol, 1.5 eq.) was added with a syringe. The reaction mixture was stirred at r. t. for 20 hours under argon atmosphere. The crude reaction mixture was diluted with Et₂O (40 ml), filtered, washed with H₂O (25 ml) and brine (25 ml), dried with MgSO₄ and concentrated *in vacuo*. The crude product was purified by flash chromatography (EtOAc – hexane, EtOAc 0% → 4% over 30 min) affording the title compound as white or light yellow solid.

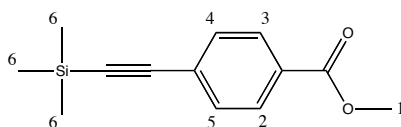
Yield = 0.455 g (98%)

Procedure 2

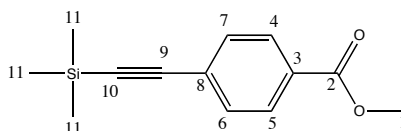
To a round bottomed flask, methyl 4-iodobenzoate (0.524 g, 2 mmol, 1 eq.), PdCl₂(PPh₃)₂ (0.042 g, 0.06 mmol, 3 mol%), CuI (0.022 g, 0.12 mmol, 6 mol%) were added, the flask was equipped with septum and flushed with argon. Then dry and degassed THF (4.2 ml) was added with a syringe. Then, a deoxygenated solution of ethynyltrimethylsilane (0.339 ml, 2.4 mmol, 1.2 eq.) in TEA (1.1 ml, 7.9 mmol, 4 eq.) was added with a syringe. The reaction mixture was stirred at 75 °C for 2 hours under argon atmosphere and then cooled down to r.t. The crude reaction mixture was diluted with Et₂O (40 ml), filtered, washed with H₂O (25 ml) and brine (25ml), dried with MgSO₄ and concentrated *in vacuo*. The crude product was purified by flash chromatography (EtOAc – hexane, EtOAc 7% over 30 min) affording the title compound as a white or light yellow solid.

Yield = 0.450 g (97%)

R_f (20 % EtOAc - hexane) = 0.48



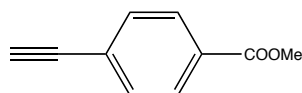
¹H NMR (300 Mhz, CDCl₃ as solvent, in ppm): δ 7.942 – 7.914 (2H, d, J = 8.4 Hz, **H2, H3**); 7.492 – 7.464 (2H, d, J = 8.4 Hz, **H5, H4**); 3.867 (3H, s, **H1**); 0.230 (9H, s, **H6**).



¹³C NMR (300 Mhz, CDCl₃ as solvent, in ppm): δ 166.80 (**C2**); 132.23 (**C6, C7**); 130.08 (**C3**); 129.75 (**C5, C4**); 128.16 (**C8**); 104.48 (**C9**); 98.02 (**C10**), 52.53 (**C1**); 0.20 (**C11**).

The compound has been previously reported ⁶⁷.

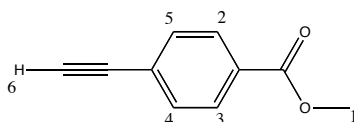
Synthesis of methyl 4-ethynylbenzoate



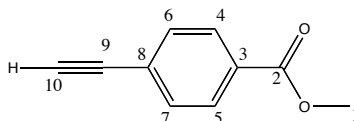
To a dry round bottomed flask, methyl 4-((trimethylsilyl)ethynyl)benzoate (0.455 g, 1.96 mmol) and K₂CO₃ (0.0414 g, 0.3 mmol) were added. Then the flask was equipped with septum and flushed with argon. Then dry and degassed MeOH (2.86 ml) and CH₂Cl₂ (1.04 ml) were added with a syringe. The reaction mixture was stirred at r.t. under Ar. – atm. for 1.5 hours, diluted with Et₂O (40 ml), filtered and concentrated *in vacuo*. The crude reaction mixture was then taken up in Et₂O (50 ml), washed with H₂O (20 ml) and brine (20 ml), dried with MgSO₄, filtered and concentrated *in vacuo*. The crude product was purified by flash chromatography (EtOAc – hexane, EtOAc 5%) affording the title compound as a white solid.

R_f (20 % EtOAc - hexane) = 0.44

Yield = 0.282 g (90%)



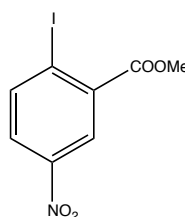
¹H NMR (300 Mhz, CDCl₃ as solvent, in ppm): δ 7.974 – 7.946 (2H, d, J = 8.4 Hz, **H2, H3**); 7.529 – 7.501 (2H, d, J = 8.4 Hz, **H4, H5**); 3.887 (3H, s, **H1**); 3.204 (1H, s, **H6**).



¹³C NMR (300 Mhz, CDCl₃ as solvent, in ppm): δ 166.34 (**C2**); 132.02 (**C6, C7**); 130.09 (**C3**); 129.41 (**C4, C5**); 126.72 (**C8**); 82.75 (**C9**); 80.01 (**C10**); 52.20 (**C1**).

The compound has been previously reported ⁶⁷.

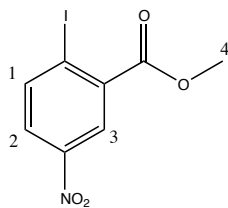
Synthesis of methyl 2-iodo-5-nitrobenzoate



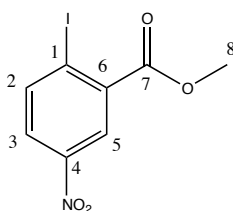
To a round bottomed flask, 2-iodobenzoic acid (2 g, 8.06 mmol) was added and dissolved in H₂SO₄ (99 %, 6 ml) at 0 °C. The mixture was stirred at 0 °C for 10 min. Then a mixture of H₂SO₄ (99 %, 4 ml) and HNO₃ (99 %, 4 ml) was added slowly with a pipette at 0 °C. The reaction flask was equipped with a condenser and the reaction mixture was stirred at 0 °C for 15 min, then warmed to r.t., stirred for 15 min, then warmed up to 130 °C and stirred at that temperature for additional 2 hours. Then the solution was cooled to 0 °C and poured into ice. The crude product was collected by suction filtration and dried under vacuum over night. Crude 2-iodo-5-nitrobenzoic acid was then dissolved in H₂SO₄ (99 %, 6.5 ml), cooled down to -5 °C and cold MeOH (20 ml) was added slowly and very carefully. The reaction mixture was stirred at 110 °C for 24 hours, cooled down to 0 °C and a large amount of water were added. The crude methyl 2-iodo-5-nitrobenzoate was obtained by extraction with CH₂Cl₂. The crude product was purified by slow recrystallisation from absolute EtOH.

Yield: 1.56 g (63%)

R_f (20% EtOAc - hexane) = 0.31



¹H NMR (300 Mhz, CDCl₃ as solvent, in ppm): δ 8.590 – 8.582 (1H, d, J = 2.4 Hz, **H3**); 8.201 – 8.172 (1H, d, J = 8.7 Hz, **H1**); 7.963 – 7.925 (1H, dd, J = 2.7 Hz, J = 8.7 Hz, **H2**); 3.968 (3H, s, **H4**).

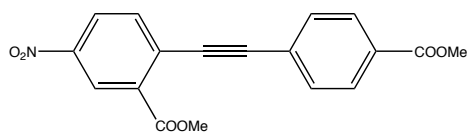


¹³C NMR (300 Mhz, CDCl₃ as solvent, in ppm): δ 164.90 (**C7**); 147.65 (**C4**); 142.78 (**C2**); 136.13 (**C6**); 126.33 (**C3**); 125.50 (**C5**); 102.55 (**C1**); 53.11 (**C8**).

MS (EI) *m/z* (relative intensity): 307 (M^+ , 100%), 276 (87.7%), 246 (11.95%), 230 (23.51%), 75 (33.90%).

HR-MS: Calculated value for C₈H₆INO₄: 306.934151. Found: 306.934160 (- 0.4 ppm)

Synthesis of methyl 2-((4-(methoxycarbonyl)phenyl)ethynyl)-5-nitrobenzoate



Procedure 1

To a round bottomed flask, methyl 2-iodo-5-nitrobenzoate (0.418 g, 1.36 mmol, 1 eq.), PdCl₂(PPh₃)₂ (0.0477 g, 0.068 mmol, 5 mol%), PPh₃ (0.0089 g, 0.034 mmol, 2.5 mol%) and CuI (0.027 g, 0.14 mmol, 10 mol%) were added, the reaction flask was equipped with a septum and flushed with argon. Then dry and degassed THF (5 ml) was added with a syringe and the mixture was stirred at r.t. under Ar. – atm. for 10 min. Then a solution of methyl 4-ethynylbenzoate (0.240 g, 1.5 mmol, 1.2 eq.) in dry and degassed THF (1 ml) and TEA (0.379 ml, 2.72 mmol, 2 eq.) were added slowly with a syringe. The reaction mixture was stirred at r.t. for 48 h. under Ar. atm. Then the reaction mixture was diluted with

EtOAc (40 ml), filtered, washed with H₂O (25 mL) and brine (25 ml), dried with MgSO₄ and concentrated *in vacuo* to afford the crude product. The crude product was purified by flash chromatography (EtOAc – hexane, EtOAc 0% - 50% over 30 min) to afford pure product as yellow solid.

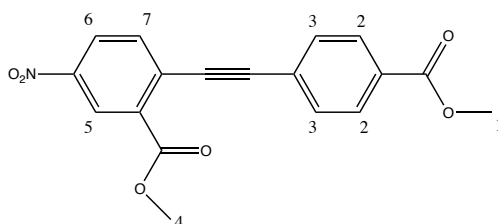
Yield: 0.373 g (81%)

Procedure 2

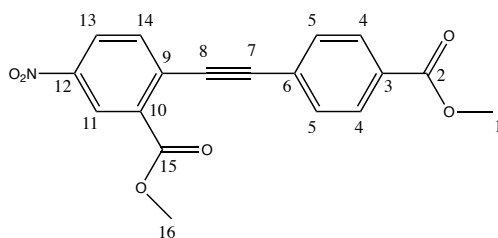
To a round bottomed flask, methyl 2-iodo-5-nitrobenzoate (0.482 g, 1.57 mmol, 1 eq.), PdCl₂(PPh₃)₂ (0.033 g, 0.047 mmol, 3 mol%) and CuI (0.018 g, 0.094 mmol, 6 mol%) were added, the reaction flask was equipped with a septum and flushed with argon. Then dry and degassed THF (7 ml) was added with a syringe and the mixture was stirred at r.t. under Ar. – atm. for 10 min. Then a deoxygenated solution of methyl 4-ethynylbenzoate (0.302 g, 1.88 mmol, 1.2 eq.) in dry and degassed THF (2 ml) and TEA (0.88 ml, 6.28 mmol, 4 eq.) were added slowly with a syringe. The reaction mixture was stirred at 58 °C for 24 h. under Ar. atm. After the completion of the reaction, the reaction mixture was allowed to cool down to r.t. Then the reaction mixture was diluted with EtOAc (40 ml), filtered, washed with H₂O (25 mL) and brine (25 ml), dried with MgSO₄ and concentrated *in vacuo* to afford the crude product. The crude product was purified by flash chromatography (EtOAc – hexane, EtOAc 0% - 50% over 30 min) to afford pure product as yellow solid.

Yield: 0.505 g (95%)

R_f (20 % EtOAc - hexane) = 0.27



¹H NMR (300 Mhz, CDCl₃ as solvent, in ppm): δ 8.836 – 8.828 (1H, d, J = 2.4 Hz, **H5**); 8.345 – 8.308 (1H, dd, J = 8.4 Hz, J = 2.4 Hz, **H6**); 8.054 – 8.026 (2H, d, J = 8.4 Hz, **H2**); 7.806 – 7.778 (1H, d, J = 8.4 Hz, **H7**); 7.662 – 7.634 (2H, d, J = 8.4 Hz, **H3**); 4.000 (3H, s, **H4**); 3.919 (3H, s, **H1**).

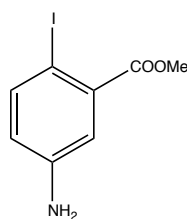


¹³C NMR (300 Mhz, CDCl₃ as solvent, in ppm): δ 166.33 (C2, C15); 164.32 (C2, C15); 146.68 (Ar-C); 135.05 (C14); 132.95 (C); 131.94 (C4, C5); 130.64 (Ar-C); 129.68 (Ar-C); 129.63 (C4, C5); 126.77 (Ar-C); 126.15 (C13); 125.85 (C11); 98.97 (C8); 89.29 (C7); 52.85 (C16); 52.35 (C1).

MS (EI) *m/z* (relative intensity): 339 (M⁺, 100%), 308 (67%), 293 (5%), 280 (11%), 262 (15%).

HR-MS: Calculated value for C₁₈H₁₃NO₆: 339.074287. Found: 339.074287 (- 0.4 ppm)

Synthesis of methyl 5-amino-2-iodobenzoate



Procedure 1

To a round bottomed flask, methyl 2-iodo-5-nitrobenzoate (0.5 g, 1.63 mmol, 1 eq.), Na₂S₂O₄ (2.84 g, 16.3 mmol, 10 eq. or 4.3 g, 24.5 mmol, 15 eq.), H₂O (10 ml) and THF (5 ml) were added. The reaction mixture was stirred for 16 hours at r. t. under N₂ – atm., diluted with EtOAc, filtered and concentrated *in vacuo*. Then it was taken up in EtOAc, washed with water and 1M HCl, dried with MgSO₄, filtered and concentrated *in vacuo*. The crude product was purified by flash chromatography (EtOAc – hexane, EtOAc 10% → 40% over 20 min) to afford pure product as yellow oil.

Yield: 0.174 g (39%)

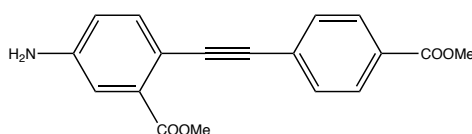
Procedure 2

To a round bottomed flask, methyl 2-iodo-5-nitrobenzoate (1.09 g, 3.5 mmol, 1 eq.), NH₄Cl (sat., aq., 3 ml), H₂O (3 ml) and THF (10 ml) were added at r.t. The mixture was stirred until all solid material was dissolved and then cooled to -5 °C. Zn dust (1.83 g, 8 eq.) was added in small portions at -5 °C. The

reaction mixture was then stirred for 2 hours at $-5\text{ }^{\circ}\text{C}$, then was let to warm up to r.t. and was stirred for 16 hours at r. t. under inert atmosphere, diluted with EtOAc, filtered and concentrated *in vacuo* in order to remove EtOAc and THF. Then it was taken up in EtOAc, washed with water and 1M HCl, dried with MgSO_4 , filtered and concentrated *in vacuo*. The crude product was purified by flash chromatography (EtOAc – hexane, EtOAc 10% \rightarrow 40% over 20 min) to afford pure product as yellow oil.

Yield: 0.866 g (90%)

Synthesis of methyl 5-amino-2-((4-(methoxycarbonyl)phenyl)ethynyl)benzoate



Procedure 1

To a round bottomed flask, methyl 2-((4-(methoxycarbonyl)phenyl)ethynyl)-5-nitrobenzoate (0.166 g, 0.491 mmol, 1 eq.), $\text{Na}_2\text{S}_2\text{O}_4$ (0.854 g, 4.91 mmol, 10 eq. or 1.28 g, 7.37 mmol, 15 eq.), H_2O (20 ml) and THF (10 ml) were added. The reaction mixture was stirred for 16 hours at r. t. under N_2 – atm., diluted with EtOAc (30 ml), filtered and concentrated *in vacuo*. Then it was taken up in EtOAc (60 ml), washed with H_2O (20 ml) and 1M HCl solution (20 ml), dried with MgSO_4 , filtered and concentrated *in vacuo*. The crude product was purified by flash chromatography (EtOAc – hexane, EtOAc 15% \rightarrow 65% over 30 min) to afford pure product as yellow solid.

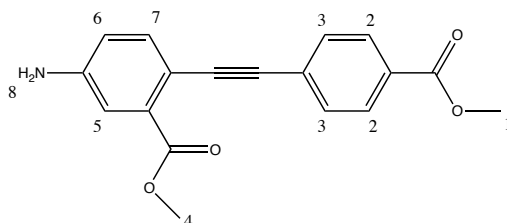
Yield: 0.062 g (41%)

Procedure 2

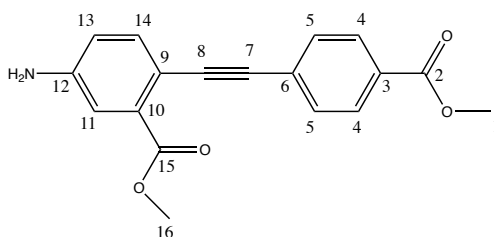
To a round bottomed flask, methyl 2-((4-(methoxycarbonyl)phenyl)ethynyl)-5-nitrobenzoate (0.179 g, 0.530 mmol, 1 eq.), NH_4Cl (sat., aq., 3 ml), H_2O (3 ml) and THF (10 ml) were added at r.t. The mixture was stirred until all solid material was dissolved and then cooled to $-5\text{ }^{\circ}\text{C}$. Zn dust (1.83 g, 8 eq.) was added in small portions at $-5\text{ }^{\circ}\text{C}$. The reaction mixture was then stirred for 2 hours at $-5\text{ }^{\circ}\text{C}$, then was let to warm up to r.t. and was stirred for 16 hours at r.t. under inert atmosphere. The crude reaction mixture was diluted with EtOAc, filtered and concentrated *in vacuo*. Then it was taken up in EtOAc, washed with water and 1M HCl, dried with MgSO_4 , filtered and concentrated *in vacuo*. The crude product was purified by flash chromatography (EtOAc – hexane, EtOAc 15% \rightarrow 65% over 30 min) to afford pure product as yellow solid.

Yield: 0.150 g (92%)

R_f (50 % EtOAc - hexane) = 0.29



¹H NMR (300 Mhz, CDCl₃ as solvent, in ppm): δ 8.000 – 7.972 (2H, d, J = 8.4 Hz, **H2**); 7.574 – 7.546 (2H, d, J = 8.4 Hz, **H3**); 7.439 – 7.411 (1H, d, J = 8.4 Hz, **H7**); 7.249 – 7.240 (1H, d, J = 2.7 Hz, **H5**); 6.781 – 6.744 (1H, dd, J = 8.4Hz, J = 2.7 Hz, **H6**); 3.988 (2H, br. s, **H8**); 3.931, 3.905 (6H, s, **H1**, **H4**).

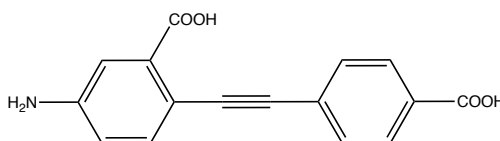


¹³C NMR (300 Mhz, CDCl₃ as solvent, in ppm): δ 166.66 (**C15**, **C2**); 146.76 (**C12**); 135.47 (**C14**); 133.18 (**Ar-C**); 131.25 (**C5**); 129.44 (**C4**); 128.97 (**Ar-C**); 128.77 (**Ar-C**); 117.87 (**C13**); 116.29 (**C11**); 112.14 (**Ar-C**); 92.12 (**C8**); 90.84 (**C7**); 52.17 (**Ar-C**); 52.14 (**C1**, **C16**).

MS (EI) *m/z* (relative intensity): 309 (M⁺, 100%), 278 (18%), 251 (6%), 235 (21.07%), 193 (7.33%).

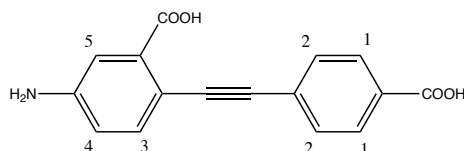
HR-MS: Calculated value for C₁₈H₁₅NO₄: 309.10011. Found: 309.10018 (+ 2.3 ppm).

Synthesis of 5-amino-2-((4-carboxyphenyl)ethynyl)benzoic acid



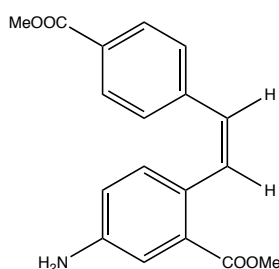
To a round bottomed flask, methyl 5-amino-2-((4-(methoxycarbonyl)phenyl)ethynyl)benzoate (0.22 g, 0.72 mmol, 1 eq), THF and H₂O (4 ml, 1:1, v:v) were added. The reaction flask was equipped with

septum and degassed with argon. Then an aqueous solution of LiOH (0.051 g, 2.16 mmol, 3 eq., 2M) was added to the reaction flask with a syringe. The reaction mixture was stirred under argon atmosphere over the night at r.t. and then acidified to pH 2 with conc. HCl. The desired product precipitated after addition of the HCl. The crude product was filtered off. ¹H NMR showed that the crude product was ca 90% pure. The desired product was isolated in ca 85% yield.



¹H NMR (300 Mhz, DMSO-d₆ as solvent, in ppm): δ 7.947 – 7.919 (2H, d, J = 8.4 Hz, **H1**, **H2**); 7.553 – 7.525 (2H, d, J = 8.4 Hz, **H1**, **H2**); 7.445 – 7.417 (1H, d, J = 8.4 Hz, **H3**); 7.291 – 7.284 (1H, d, J = 2.1 Hz, **H5**); 6.948 – 6.912 (1H, d, J = 8.4 Hz, J = 2.1 Hz, **H4**).

Attempted synthesis of (*Z*)-methyl 5-amino-2-(4-(methoxycarbonyl)styryl)benzoate



Procedure 1. Using the Pd(OAc)₂/KOH/DMF–system

For the amount of different reagents, reaction temperatures etc., please see section 10.4.9, table 8.

To a round bottomed flask, methyl 5-amino-2-((4-(methoxycarbonyl)phenyl)ethynyl)benzoate (1 eq), Pd(OAc)₂ (2 - 6 mol%), KOH (1.5 eq) and degassed DMF (1,3,1 or 1 ml pr. 0.5 mmol substrate) were added. The reaction flask was equipped with a condenser and flushed with argon for 20 min. The reaction mixture was stirred for 6 h. at 145 °C or 12 h. at 100 °C under Ar. atm., cooled to r.t. The crude reaction mixture was washed with EtOAc and filtered. The organic layer was washed with H₂O (5 times) and brine (3 times), dried with MgSO₄ and concentrated *in vacuo*. The crude product was purified twice by flash chromatography (EtOAc – hexane, EtOAc 25 – 55% over 45 min) affording the desired product in ca 25% yield, ca 65% pure based on NMR analysis.

Procedure 2. Using the Pd(OAc)₂/PPh₃/NaOMe/THF–system

For the amount of different reagents, reaction temperatures etc., please see section 10.4.10, table 9.

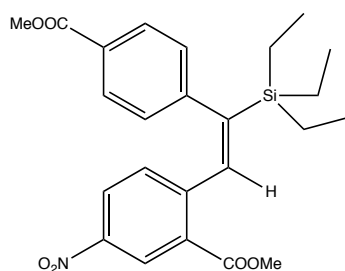
To a round bottomed flask, methyl 5-amino-2-((4-(methoxycarbonyl)phenyl)ethynyl)benzoate (1 eq.), Pd(OAc)₂ (5 or 10 mol%), PPh₃ (5 or 10 mol%) and NaOMe (5 or 10 eq.) were added. The reaction flask was equipped with septum and flushed with argon for 15 min. Then dry and degassed THF was added with a syringe and the reaction mixture was stirred under Ar. atm. at r.t. for 12 or 48 h. or at 45 °C for 12 h.. The crude reaction mixture was taken up in EtOAc, filtered, washed with H₂O and brine, dried with MgSO₄ and concentrated *in vacuo* giving the crude product. The crude product was purified by flash chromatography (EtOAc – hexane, EtOAc 25 – 55% over 45 min) affording only the starting material in ca 85 – 90 % yield.

Procedure 3. Using the hydroboration/protonolysis approach

For the amount of different reagents, reaction temperatures etc., please see section 10.4.12, table 10.

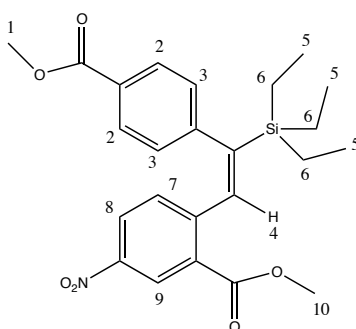
To a round bottomed flask, cyclohexene (4, 8, 20, 30 or 40 eq.) and THF was added. The flask was equipped with septum, degassed with argon and cooled down to -20 °C. Then degassed solution of BH₃•THF complex (2, 4, 10, 15 or 20 eq.) was added slowly with a syringe. The mixture was stirred at -20 °C for 10 min, then was allowed to warm slowly to -5 °C. Then it was stirred for 1 h. at -5 °C, during that time white solid was formed. Then the mixture was cooled to -10 °C. Pre-cooled solution (0 °C) of methyl 5-amino-2-((4-(methoxycarbonyl)phenyl)ethynyl)benzoate (1 eq.) in THF was added to the solution of dicyclohexylborane at -10 °C, stirred for 10 min., then warmed to -5 °C and stirred for 1-2 hours under Ar. atm.. Then anhydrous AcOH (12, 24, 60, 80 or 120 eq.) was added slowly with a syringe at -5 °C and the reaction mixture was stirred at -5 °C for 0.5-3 h. The crude reaction mixture was taken up in EtOAc, filtered, washed with H₂O and brine, dried with MgSO₄ and concentrated *in vacuo* giving the crude product. The crude product was purified by flash chromatography (EtOAc – hexane, EtOAc 25 – 55% over 45 min) affording only the starting material in ca 90 – 95 % yield.

Synthesis of (*E*)-methyl 2-(2-(4-(methoxycarbonyl)phenyl)-2-(triethylsilyl)vinyl)-5-nitrobenzoate



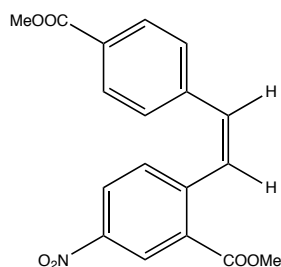
For the amount of different reagents, reaction temperatures etc., please see section 10.4.13, table 11.

To a round bottomed flask, methyl 2-((4-(methoxycarbonyl)phenyl)ethynyl)-5-nitrobenzoate (1 eq.) and $\text{PtO}_2 \cdot \text{H}_2\text{O}$ (5 – 20 mol%) were added. The flask was equipped with septum and degassed with argon. Then, dry and degassed THF was added with a syringe. The reaction mixture was stirred under r.t. for 15 min. and then degassed Et_3SiH (1.5 – 10 eq.) was added with a syringe. Inert gas inlet/outlet was removed and the mixture was stirred at 55 – 70 °C for 2 – 3.5 h. After this, the reaction mixture was cooled down to r.t., taken up in small volume of EtOAc and filtered. The filtrated mixture was concentrated *in vacuo*. The crude product was purified by flash chromatography (EtOAc – hexane, EtOAc 10 – 35% over 55 min) sometimes affording the desired product as yellow oil in up to 80 % yield.



^1H NMR (300 Mhz, CDCl_3 as solvent, in ppm): δ 8.746 – 8.738 (1H, d, $J = 2.4$ Hz, **H9**); 8.216 – 8.179 (1H, dd, $J = 8.4$ Hz, $J = 2.4$ Hz, **H8**); 7.751 – 7.723 (2H, d, $J = 5.6$ Hz, **H2**); 7.092 – 7.064 (1H, d, $J = 8.4$ Hz, **H7**); 6.919 – 6.891 (2H, $J = 8.4$ Hz, **H3**); 6.867 (1H, s, **H4**), 3.808 (3H, s, **H1**, **H10**); 3.783 (3H, s, **H1**, **H10**); **H5** and **H6** are buried somewhere between 1.246 – 0.546 ppm.

Attempted synthesis of (*Z*)-methyl 2-(4-(methoxycarbonyl)styryl)-5-nitrobenzoate from (*E*)-methyl 2-(2-(4-(methoxycarbonyl)phenyl)-2-(triethylsilyl)vinyl)-5-nitrobenzoate

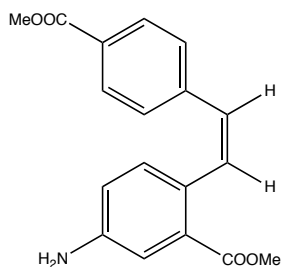


For the amount of different reagents, reaction temperatures etc., please see section 10.4.13, table 12.

To a round bottomed flask, (*E*)-methyl 2-(2-(4-(methoxycarbonyl)phenyl)-2-(triethylsilyl)vinyl)-5-nitrobenzoate, TBAF (1.1 – 3.2 eq.) and CuI (0 – 1.5 eq.) were added. The reaction flask was equipped with septum and degassed. Then dry and degassed THF was added with a syringe. The reaction mixture

was stirred at the particular temperatures for a particular amount of time (scheme X). Then the crude reaction mixture was taken up in small volume of EtOAc and filtered. The filtrated mixture was concentrated *in vacuo*. The crude product was purified by flash chromatography (EtOAc – hexane, EtOAc 10 – 35% over 35 min).

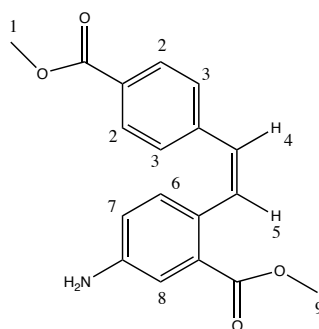
Synthesis of (*Z*)-methyl 5-amino-2-(4-(methoxycarbonyl)styryl)benzoate by reduction with Lindlar catalyst



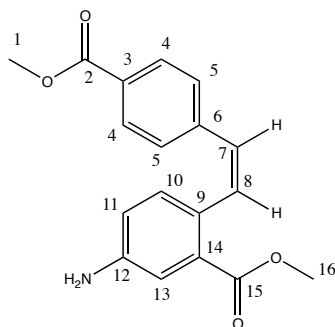
For the amount of different reagents, reaction temperatures etc., please see section 10.4.14, table 13.

To a round bottomed flask, methyl 2-((4-(methoxycarbonyl)phenyl)ethynyl)-5-nitrobenzoate (1eq.), Lindlar catalyst (5mol%), quinoline (5 – 30 mol%) and THF were added. The reaction flask was sealed with a septum and hydrogen gas was introduced into the reaction flask from a balloon via hypodermic needle that was stuck through septum. The reaction flask was flooded with hydrogen. Hydrogen gas was also bubbled through the solvent - after that the gas inlet was removed and the flask was sealed with plastic tape. The reaction mixture was stirred for 16 h. Then the crude reaction mixture was taken up in small volume of EtOAc and filtered. The filtrated mixture was concentrated *in vacuo*. The crude product was purified by flash chromatography (EtOAc – hexane, EtOAc 20 – 55% over 35 min). The desired product was once isolated as the only product of the reaction in 96 % yield. The desired product is yellow oil.

R_f (50 % EtOAc - hexane) = 0.25



¹H NMR (400 Mhz, CDCl₃ as solvent, in ppm): δ 7.801 – 7.780 (2H, d, J = 8.4 Hz, **H2, H3**); 7.292 – 7.286 (1H, d, J = 2.4 Hz, **H8**); 7.158 – 7.137 (2H, d, J = 8.4 Hz, **H2, H3**); 7.073 – 7.043 (1H, d, J = 12 Hz, **H4, H5**); 6.928 – 6.907 (1H, d, J = 8.4 Hz, **H6**); 6.589 – 6.562 (1H, dd, J = 2.4 Hz, **H7**); 6.543 – 6.513 (1H, dd, J = 12 Hz, **H4, H5**); 3.851 (6H, s, **H1, H9**).



¹³C NMR (400 Mhz, CDCl₃ as solvent, in ppm): δ 167.44 (**C2, C15**); 166.96 (**C2, C15**); 145.74 (**Ar-C**); 142.19 (**Ar-C**); 132.78 (**C7, C8**); 132.09 (**C10**); 129.99 (**Ar-C**); 129.29 (**C4, C5**); 129.07 (**C4, C5**); 128.82 (**Ar-C**); 128.02 (**Ar-C**); 127.33 (**C7, C8**); 118.67 (**C11**); 116.48 (**C13**); 52.00 (**C1, C16**); 51.95 (**C1, C16**).

MS (EI) *m/z* (relative intensity): 311 (**M⁺**, 100%), 280 (15.9%), 251 (10.60%), 220 (6%), 164 (35.03%).

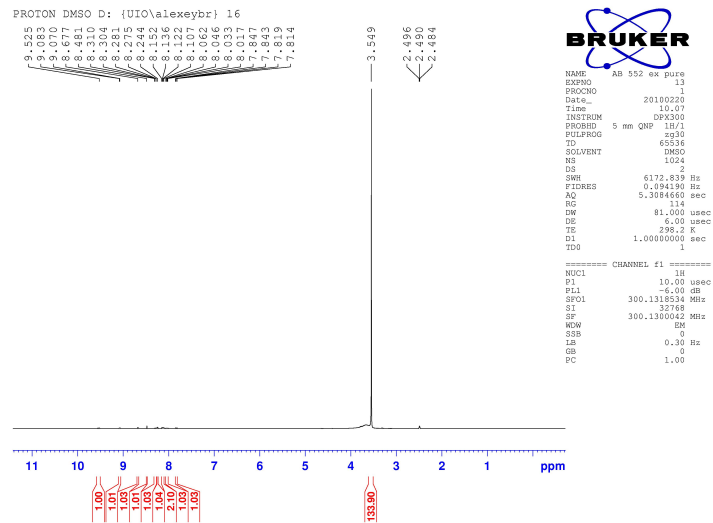
HR-MS: Calculated value for C₁₈H₁₇NO₄: 311.11576. Found: 311.115758 (- 2.2 ppm).

Attachment 1

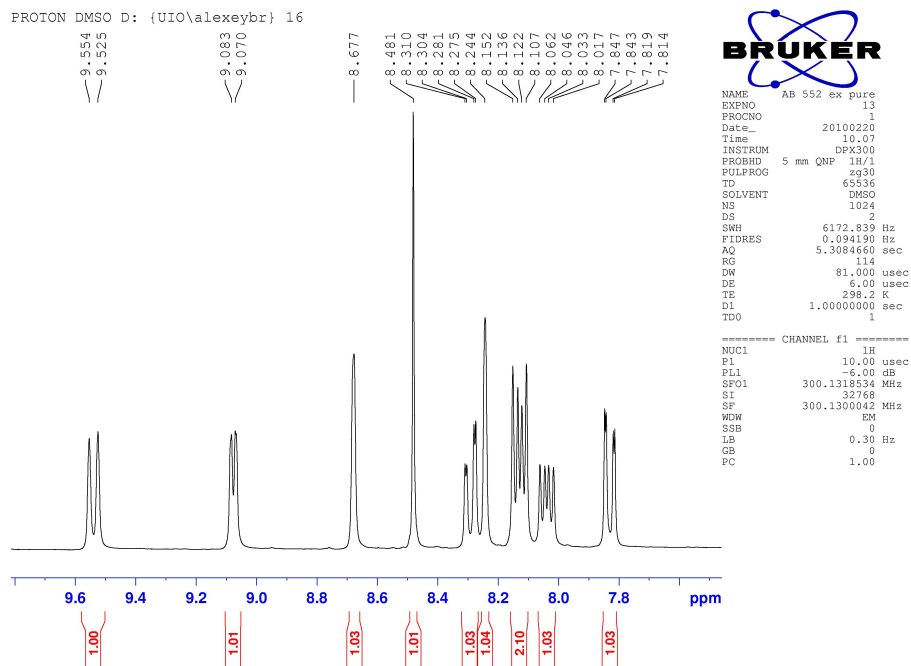
NMR data

AB 81 - (E)-8-hydroxy-7-((6-sulfonaphthalen-2-yl)diazenyl)quinoline-5-sulfonic acid

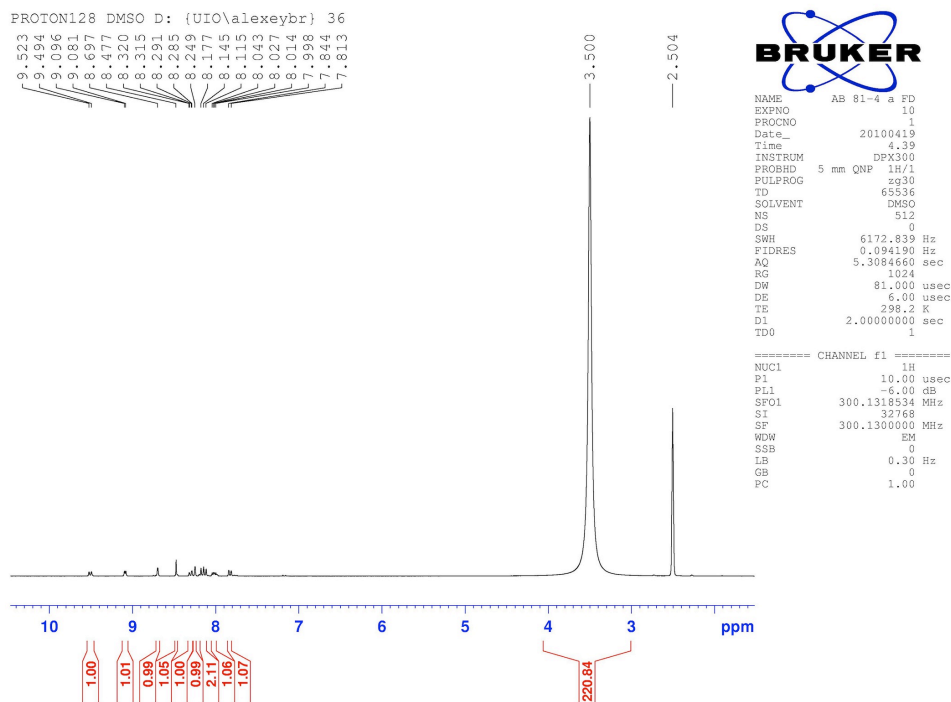
¹H NMR



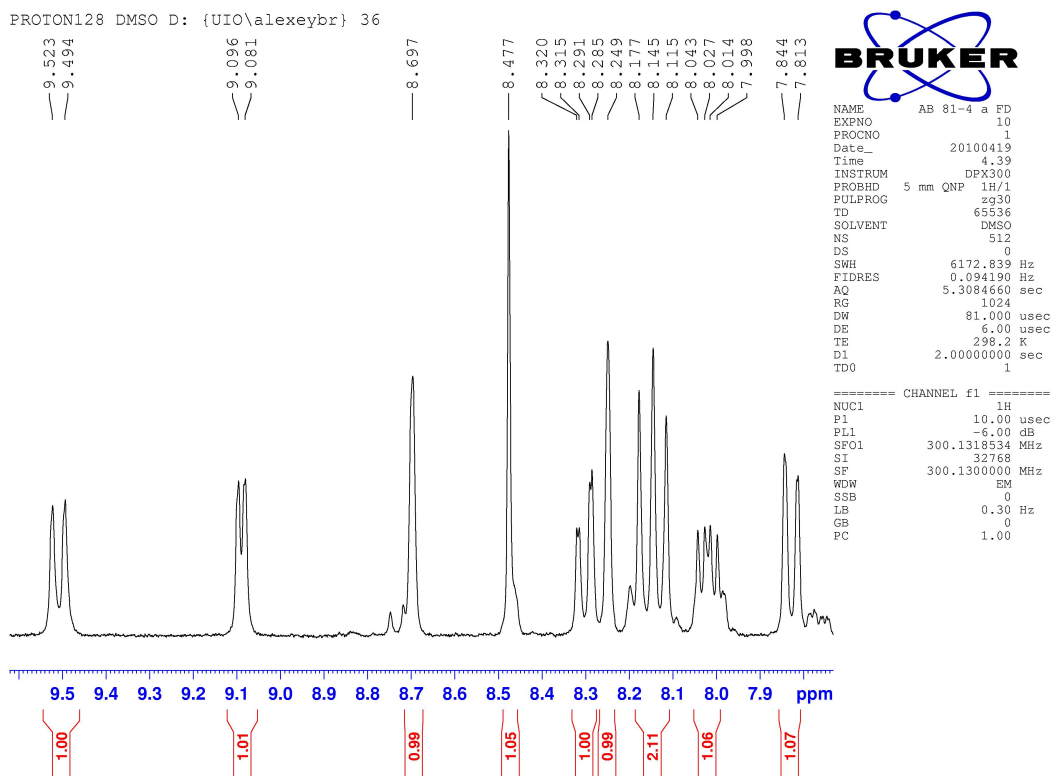
Expansion of the aromatic region



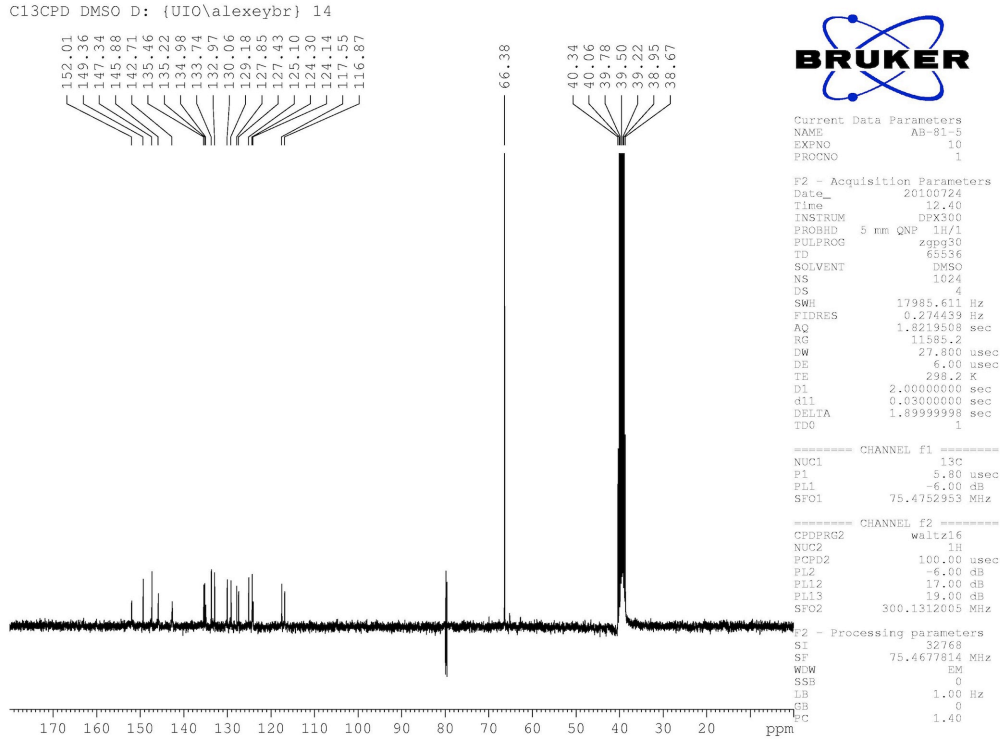
¹H NMR spectrum after freeze drying (no dioxane used during the purification)



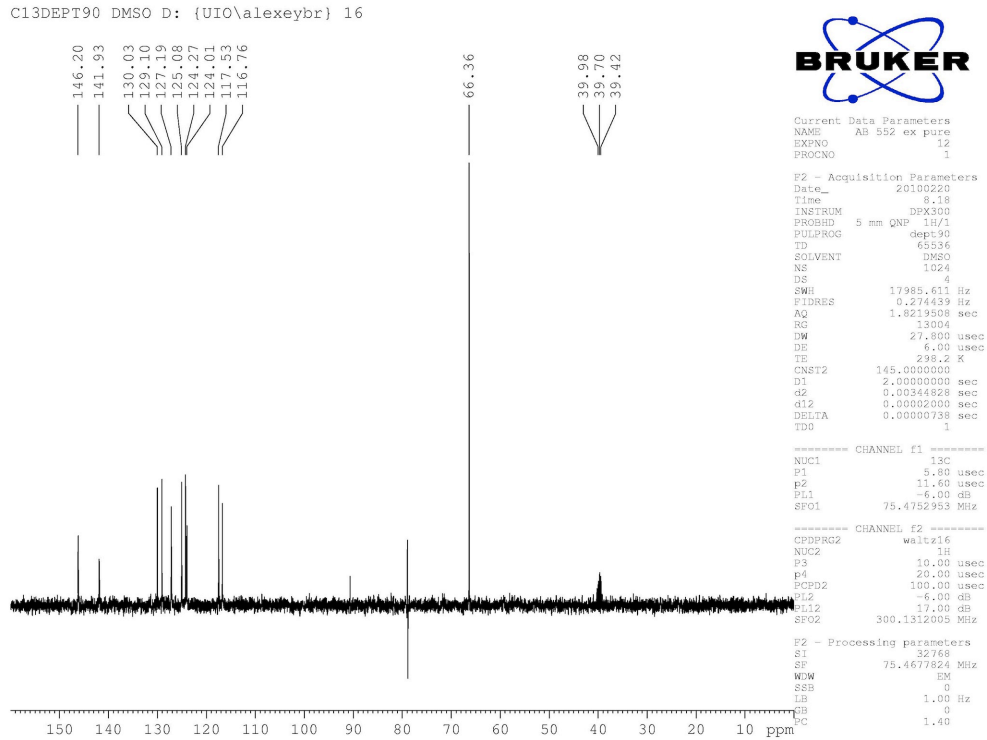
Expansion of the aromatic region



C13CPD

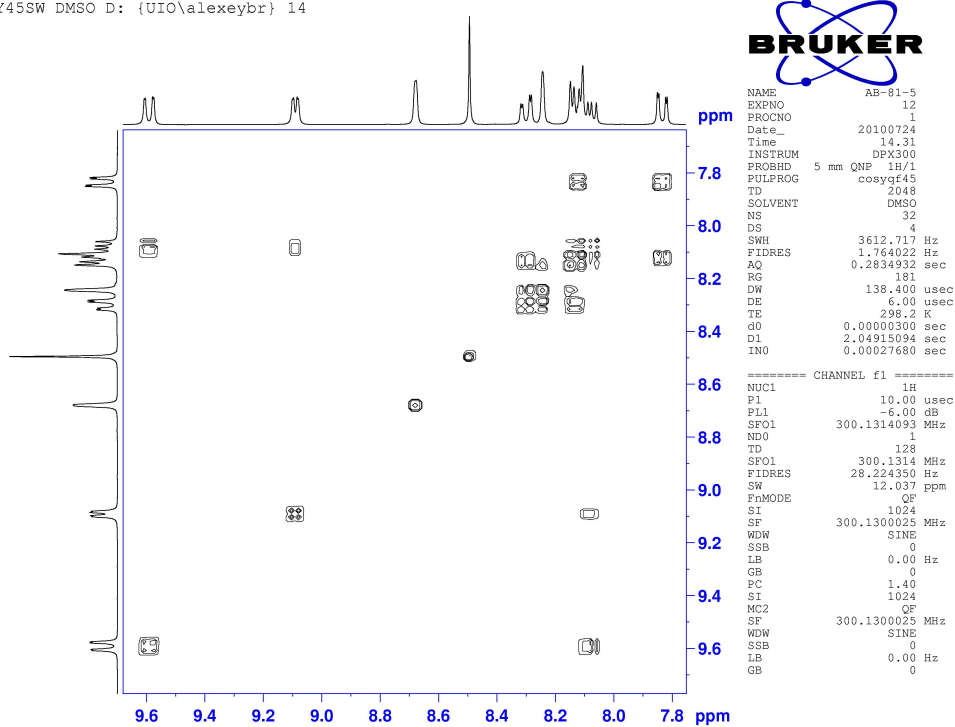


C13 DEPT90



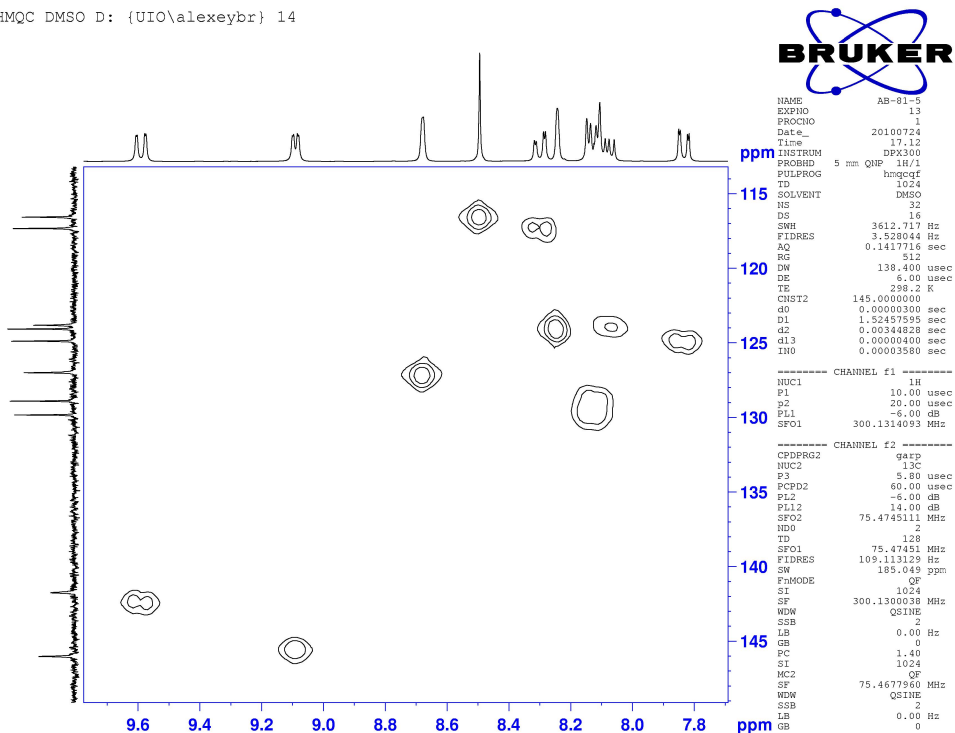
COSY45SW, expansion of the aromatic region

COSY45SW DMSO D: {UIO\alexeybr} 14



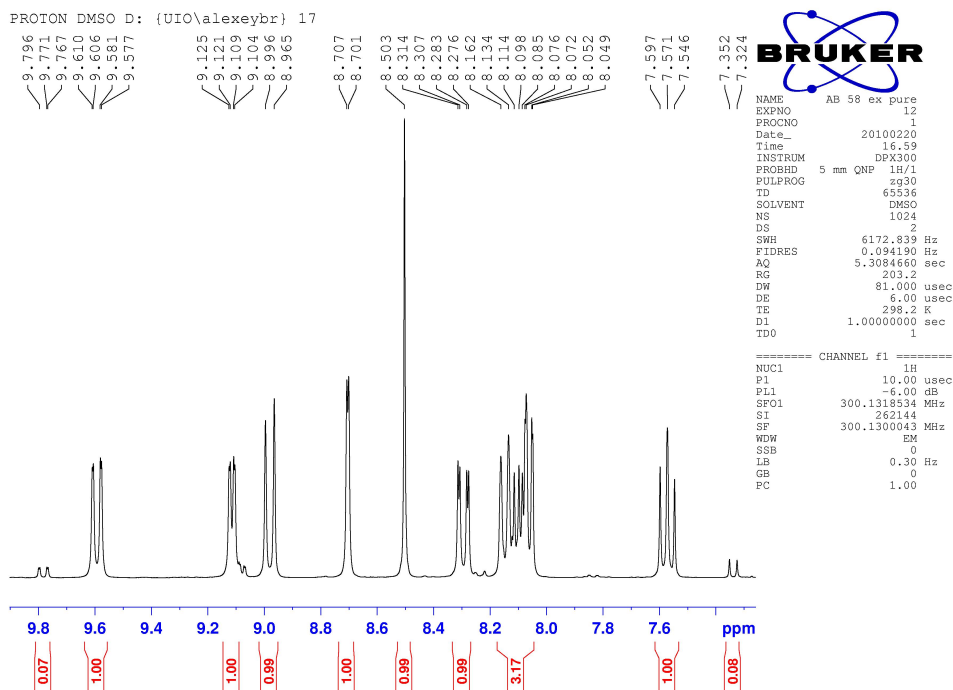
HMQC (¹H – ¹³C), expansion of the aromatic region

HMQC DMSO D: {UIO\alexeybr} 14

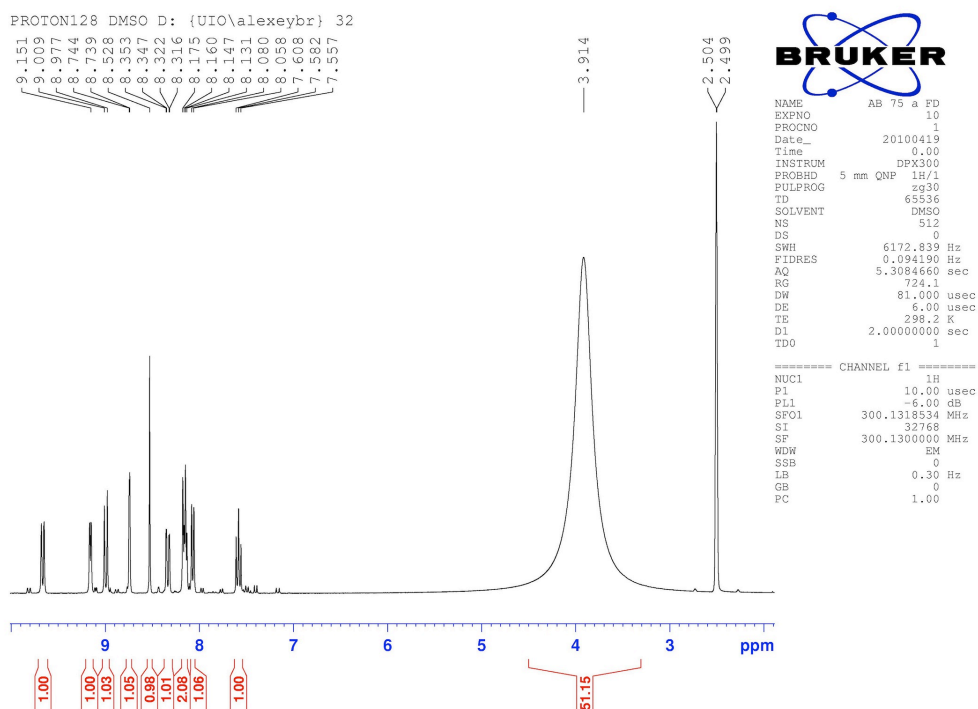


AB 75 - (E)-8-hydroxy-7-((5-sulfonaphthalen-2-yl)diazenyl)quinoline-5-sulfonic acid

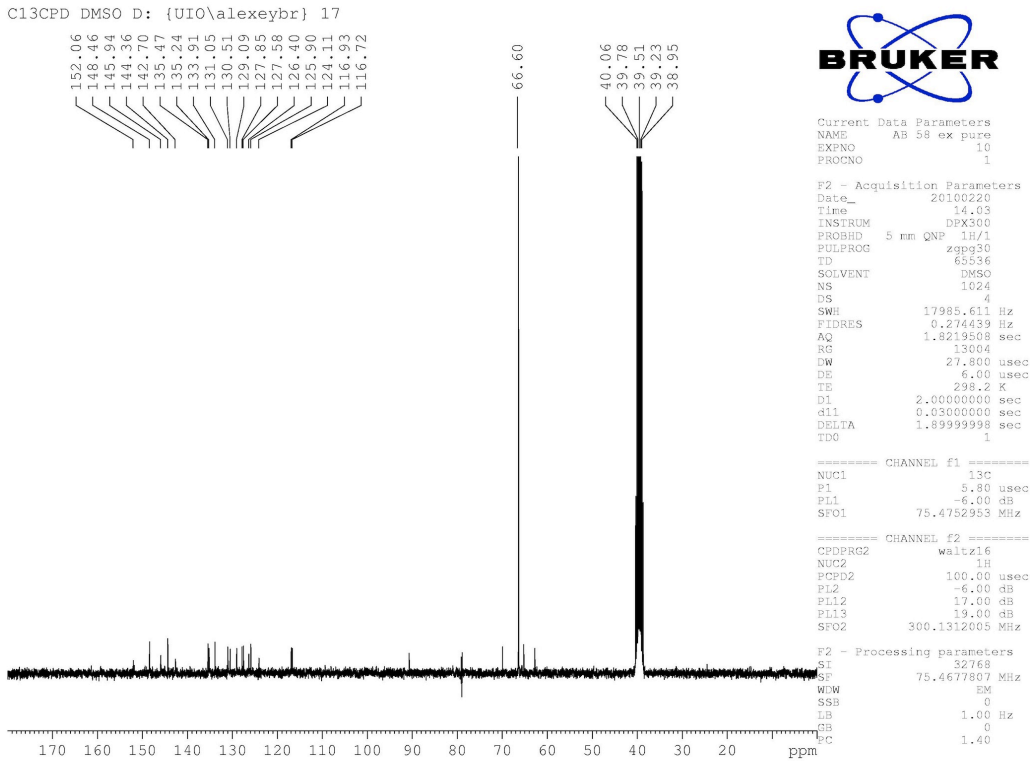
¹H NMR, expansion of the aromatic region



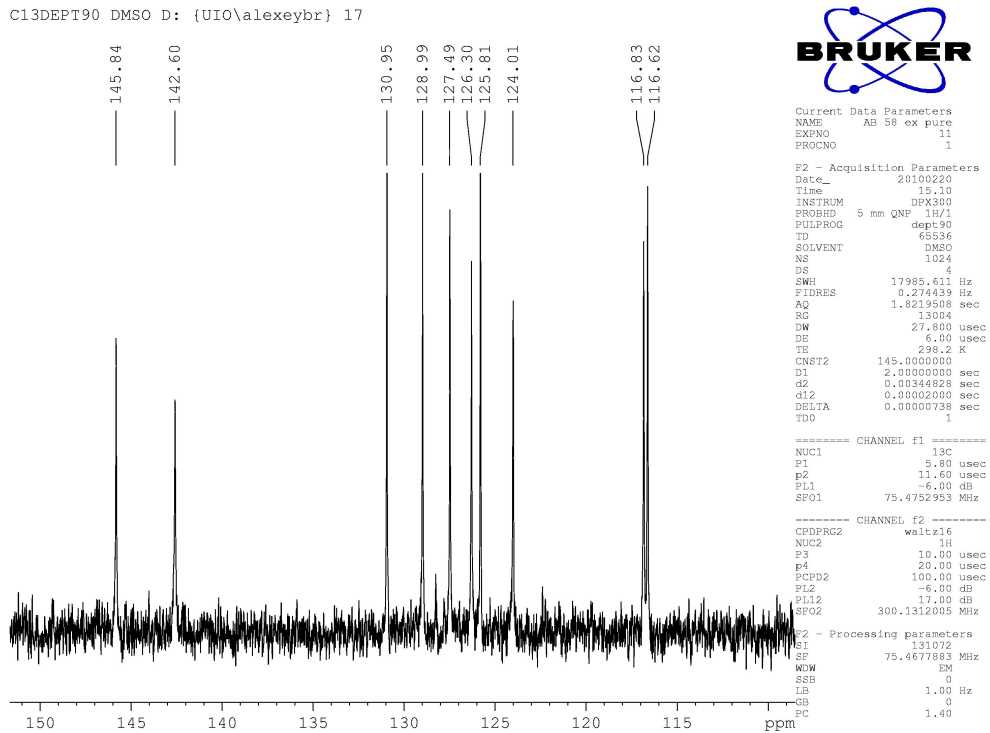
¹H NMR spectrum after freeze drying (no dioxane used during the purification)



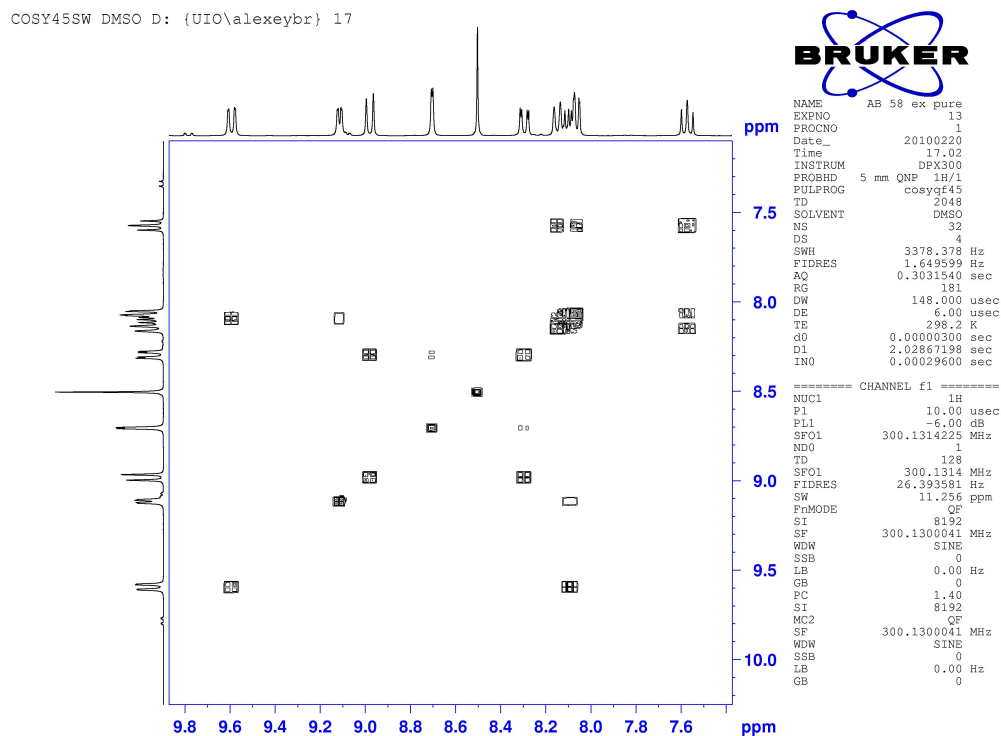
C13CPD



C13 DEPT90

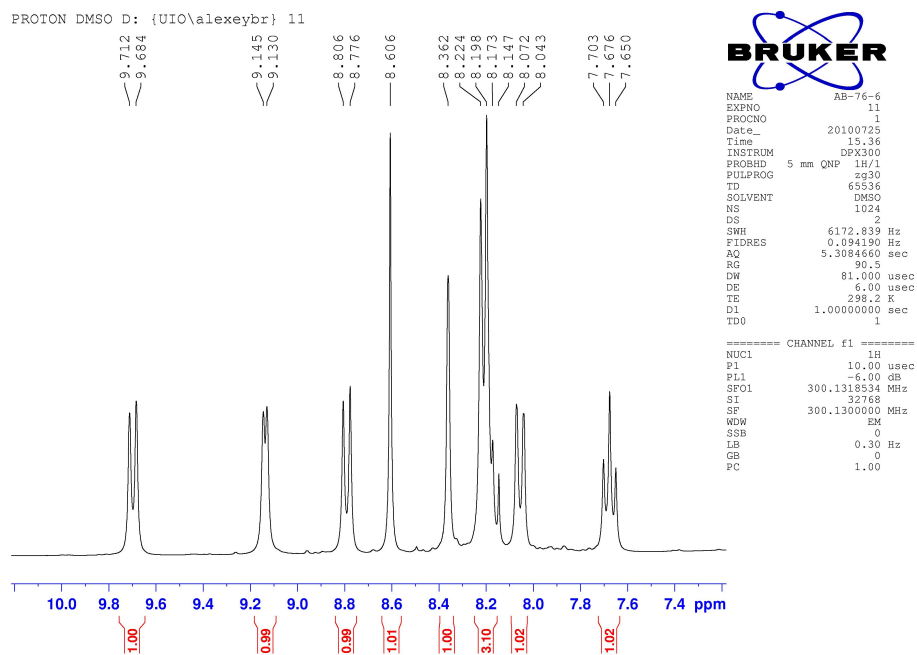


COSY45SW, expansion of the aromatic region

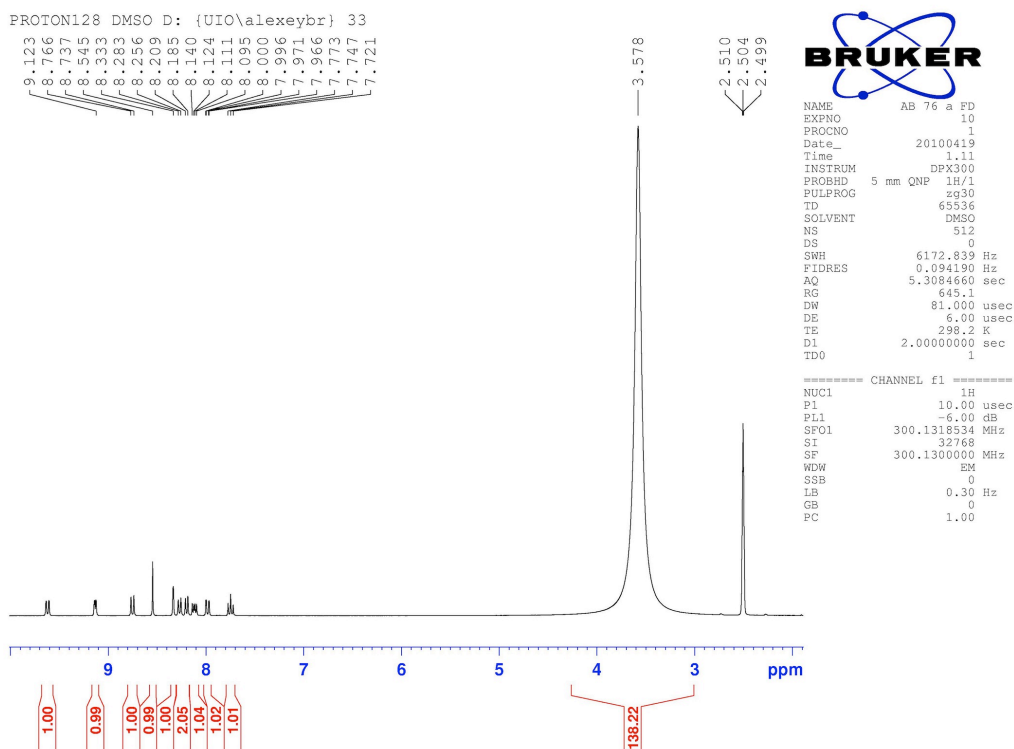


AB 76 - (E)-8-hydroxy-7-((6-sulfonaphthalen-1-yl)diazanyl)quinoline-5-sulfonic acid

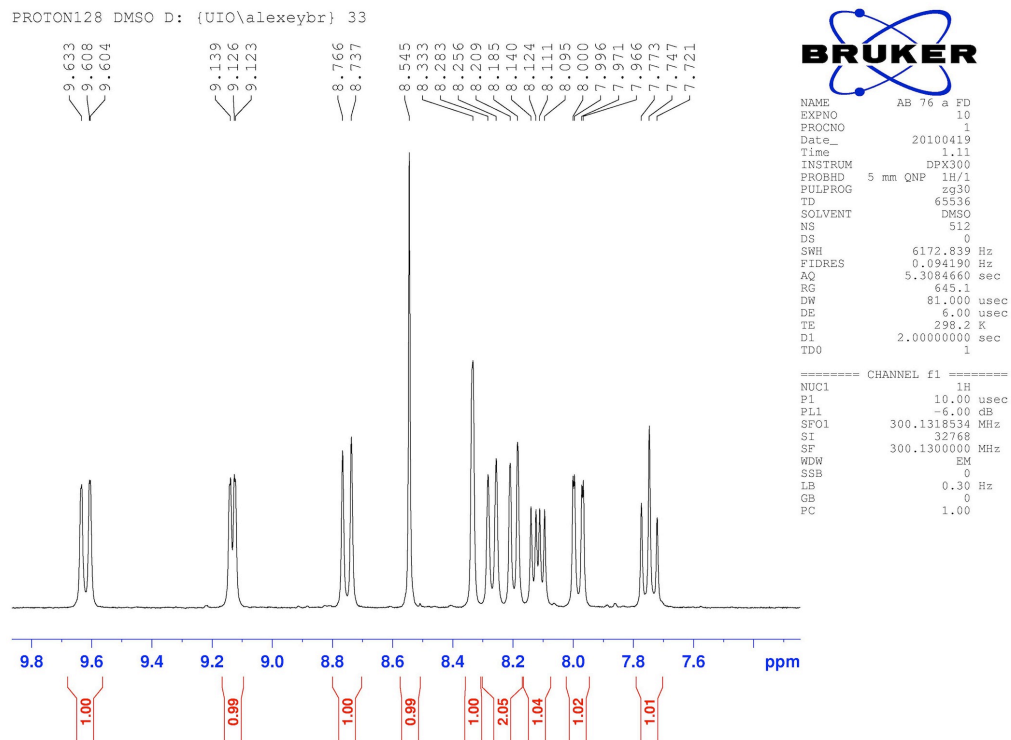
¹H NMR, expansion of the aromatic region



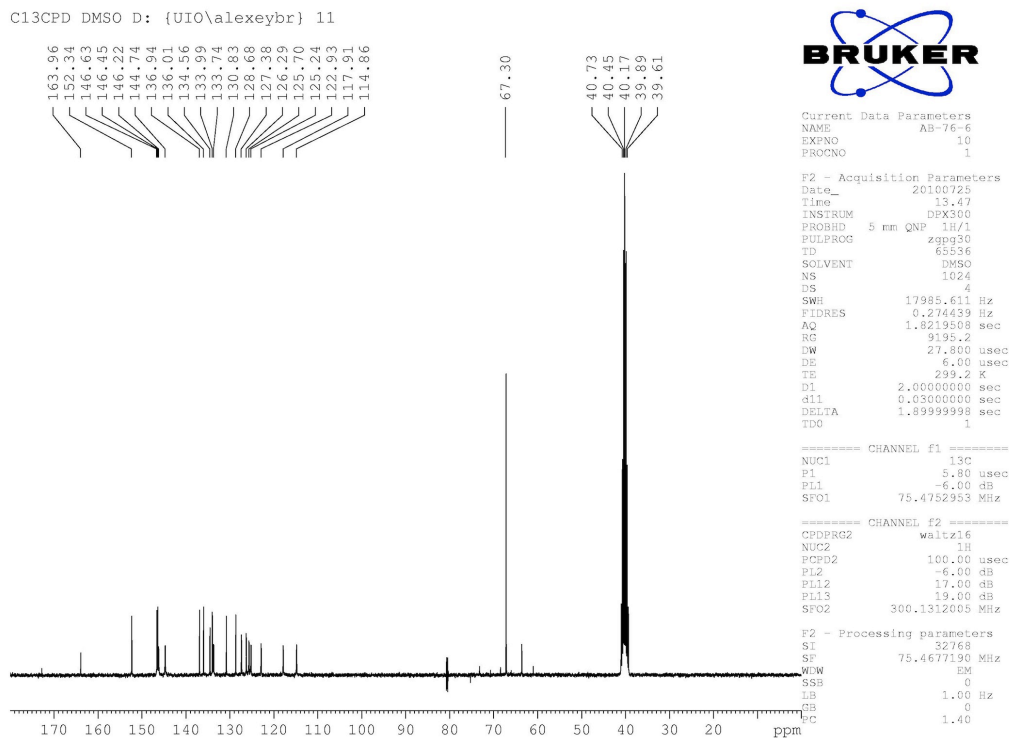
¹H NMR spectrum after freeze drying (no dioxane used during the purification)



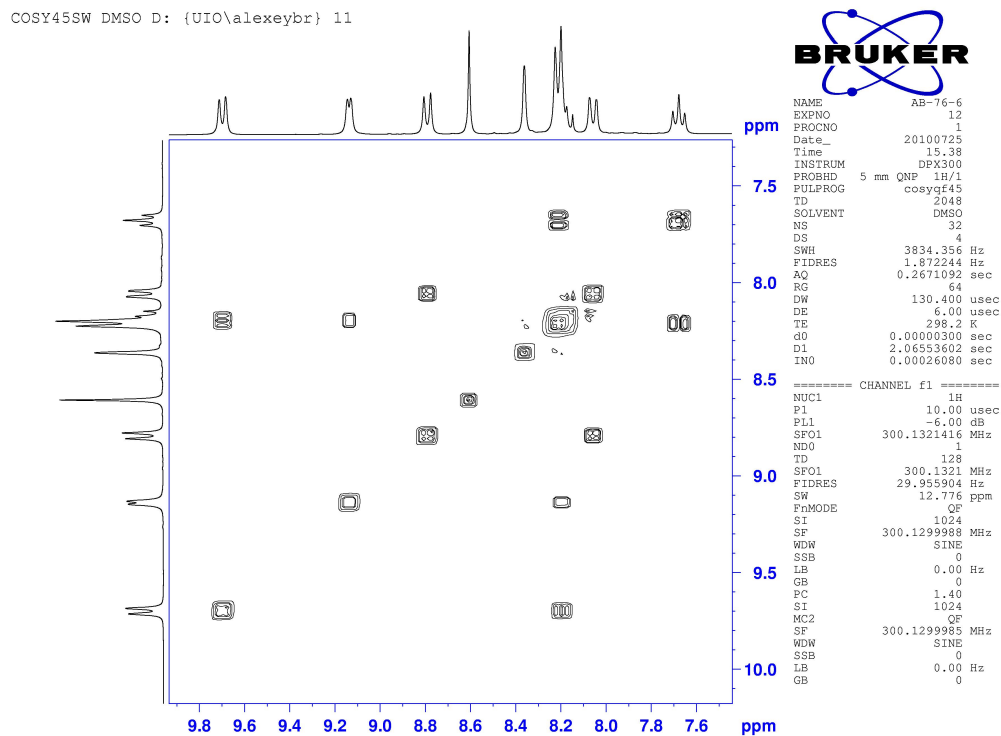
Expansion of the aromatic region



C13CPD

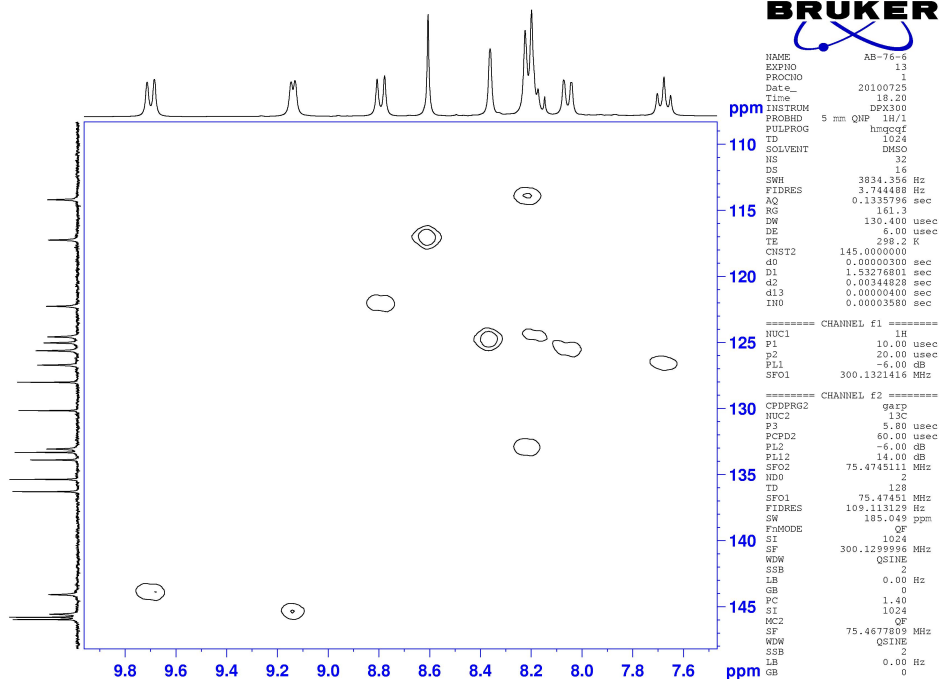


COSY45SW, expansion of the aromatic region



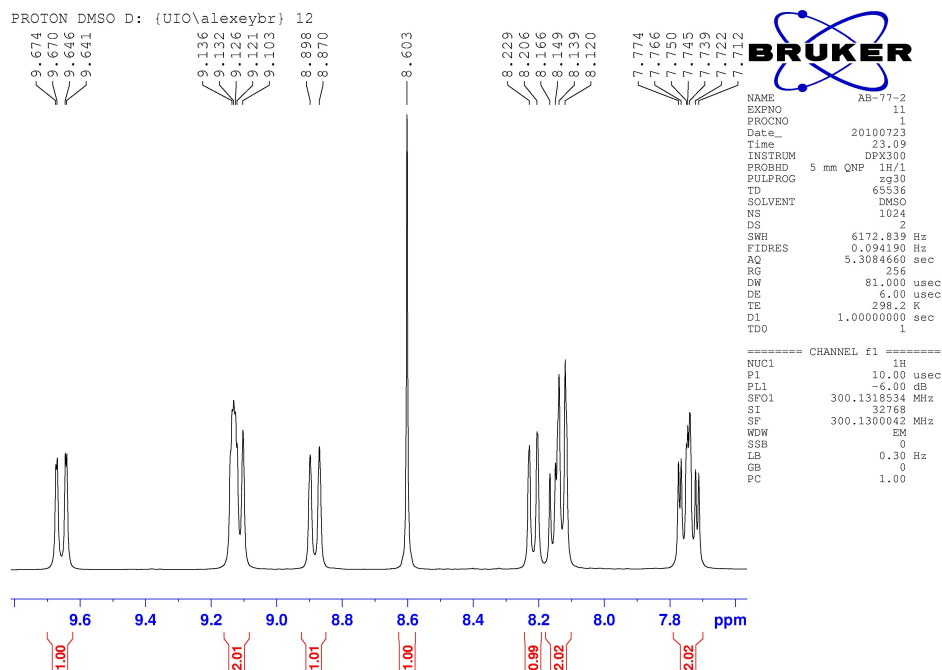
HMQC (¹H – ¹³C), expansion of the aromatic region

HMQC DMSO D: {UIO\alexeybr} 11

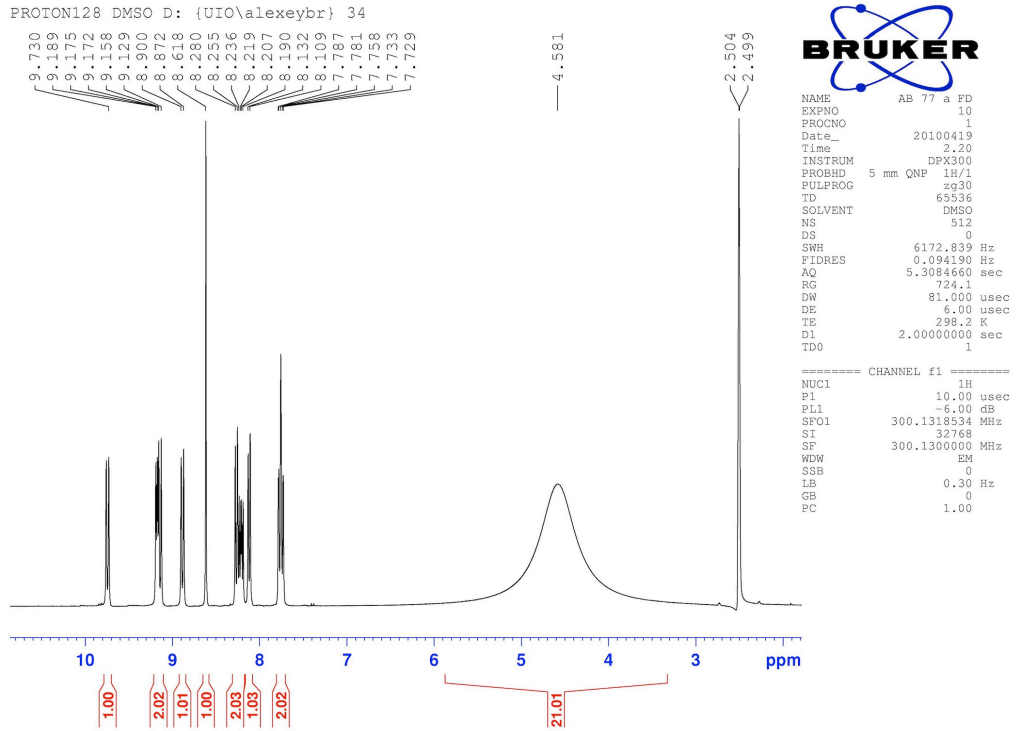


AB 77 - (E)-8-hydroxy-7-((5-sulfonaphthalen-1-yl)diazenyl)quinoline-5-sulfonic acid

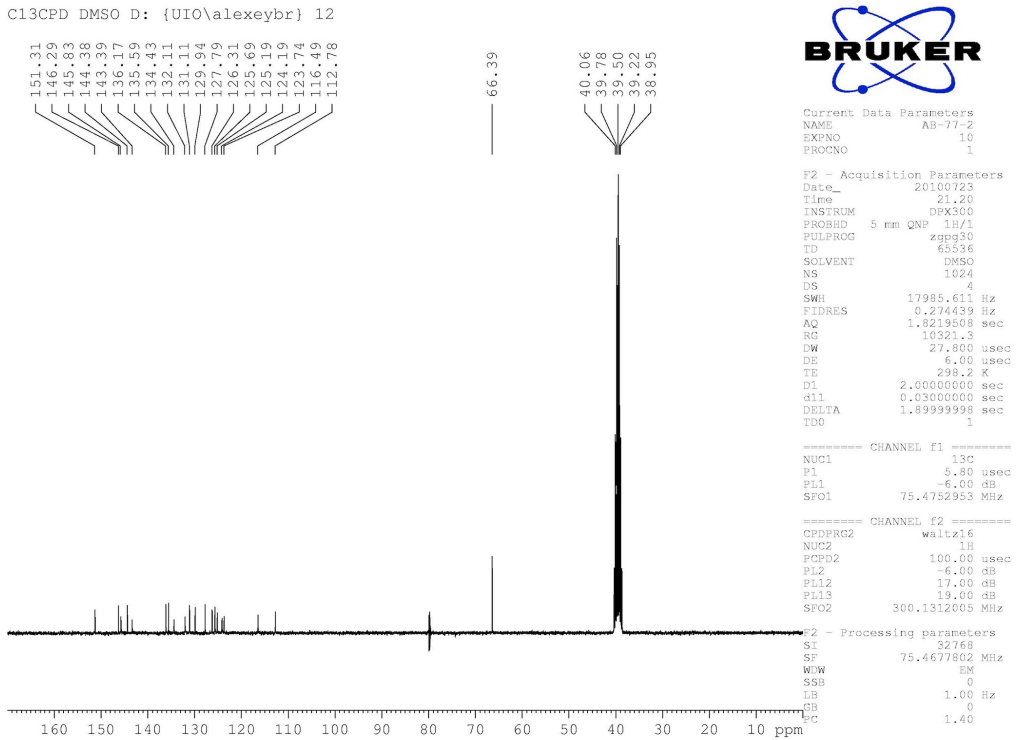
¹H NMR, expansion of the aromatic region



¹H NMR spectrum after freeze drying (no dioxane used during the purification)

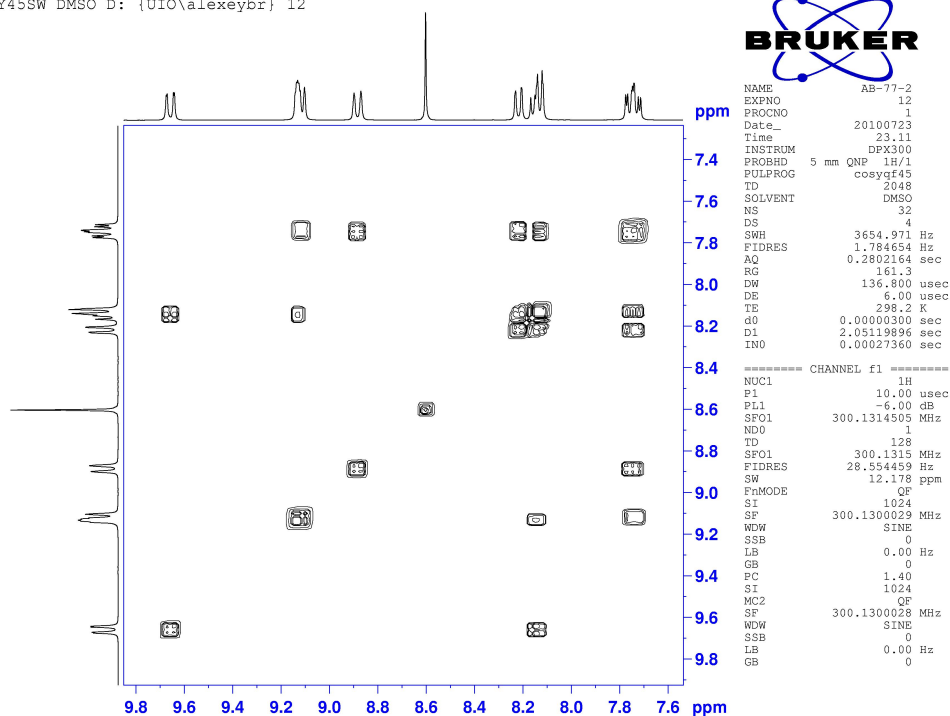


C13CPD



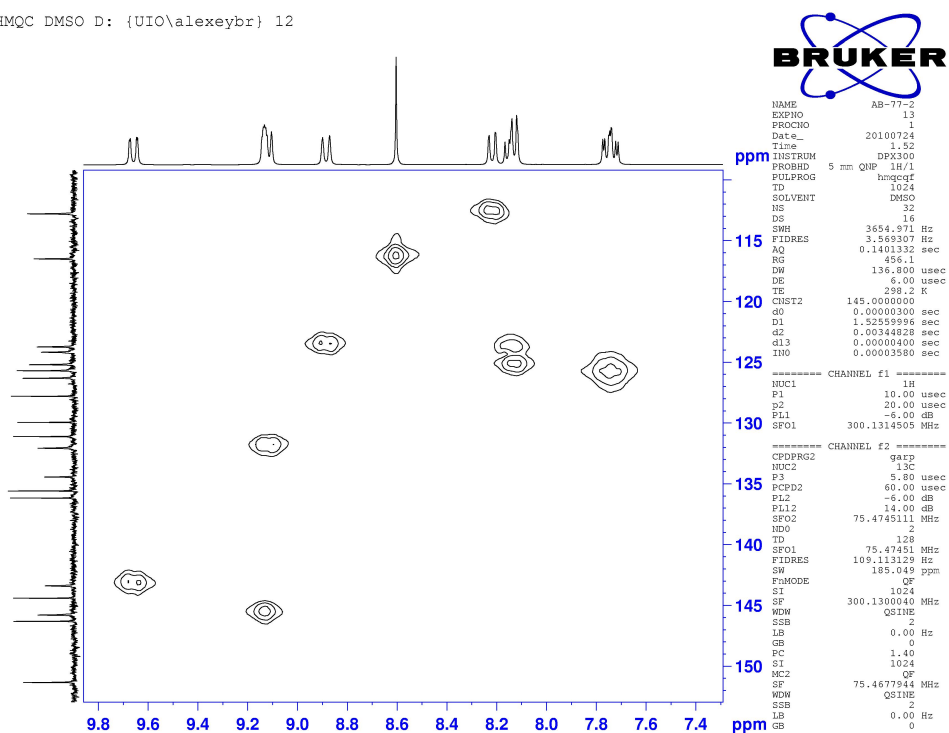
COSY45SW, expansion of the aromatic region

COSY45SW DMSO D: {UIO\alexeybr} 12



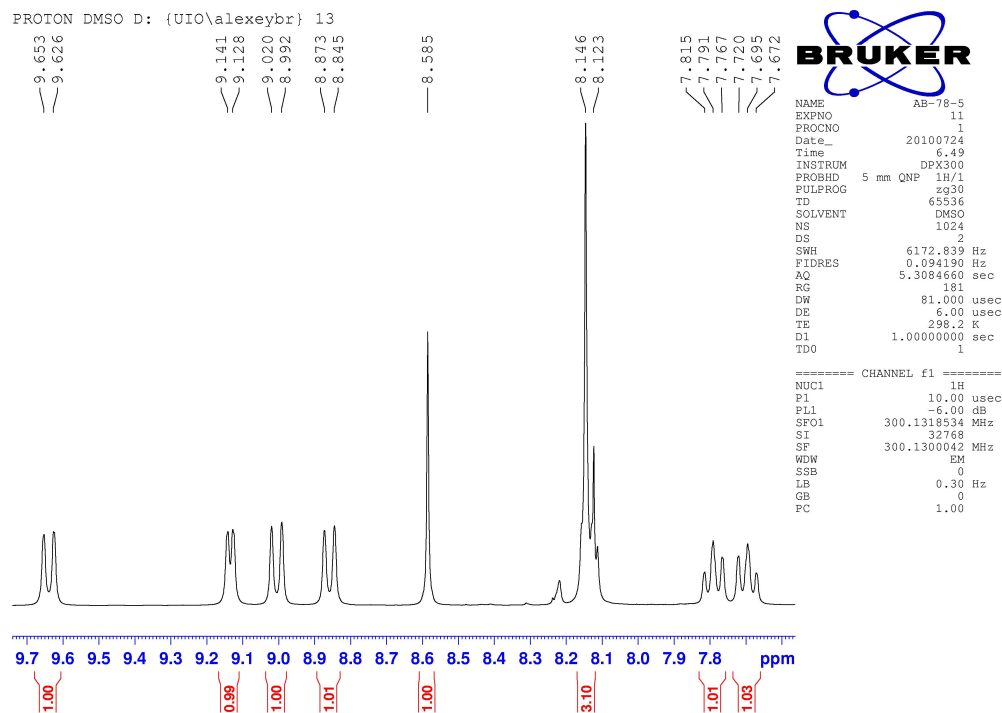
HMQC (¹H - ¹³C), expansion of the aromatic region

HMQC DMSO D: {UIO\alexeybr} 12

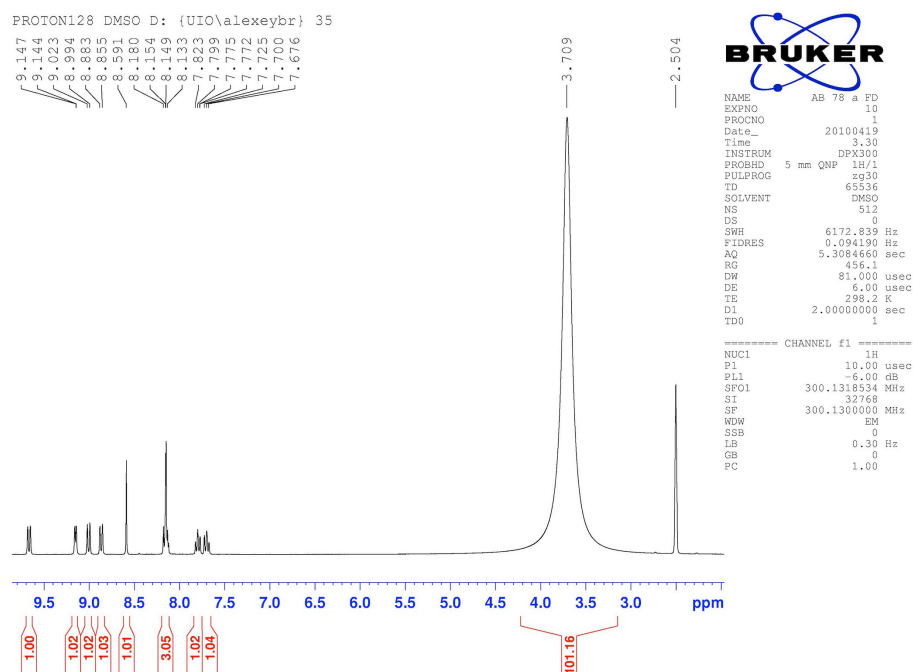


AB 78 - (E)-8-hydroxy-7-((4-sulfonaphthalen-1-yl)diazenyl)quinoline-5-sulfonic acid

¹H NMR, expansion of the aromatic region

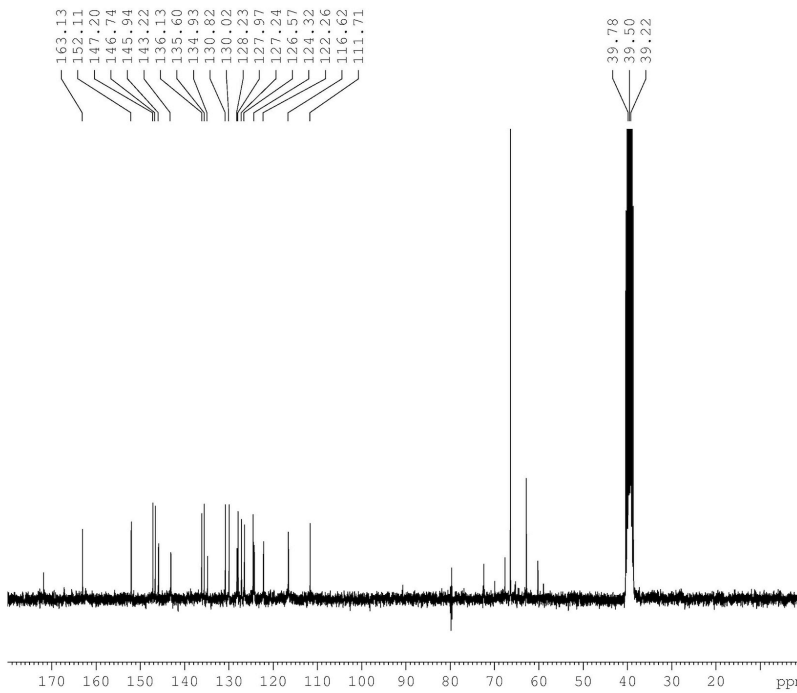


¹H NMR spectrum after freeze drying (no dioxane used during the purification)



C13CPD

C13CPD DMSO D: {UIO\alexeybr} 13



```

Current Data Parameters
NAME      AB-78-5
EXPNO     10
PROCNO    1

F2 - Acquisition Parameters
Date_     20100724
Time      5.00
INSTRUM   DFX500
PROBHD    5 mm QNP 1H/1
PULPROG   zgpg30
TD         65536
SOLVENT   DMSO
NS         1024
DS         4
SWH        17985.611 Hz
FIDRES     0.274439 Hz
AQ         1.8213508 sec
RG         11585.2
DW         27.800 usec
DE         6.00 usec
TE         298.2 K
D1         2.0000000 sec
d11        0.0300000 sec
DELTA     1.89999998 sec
TD0        1

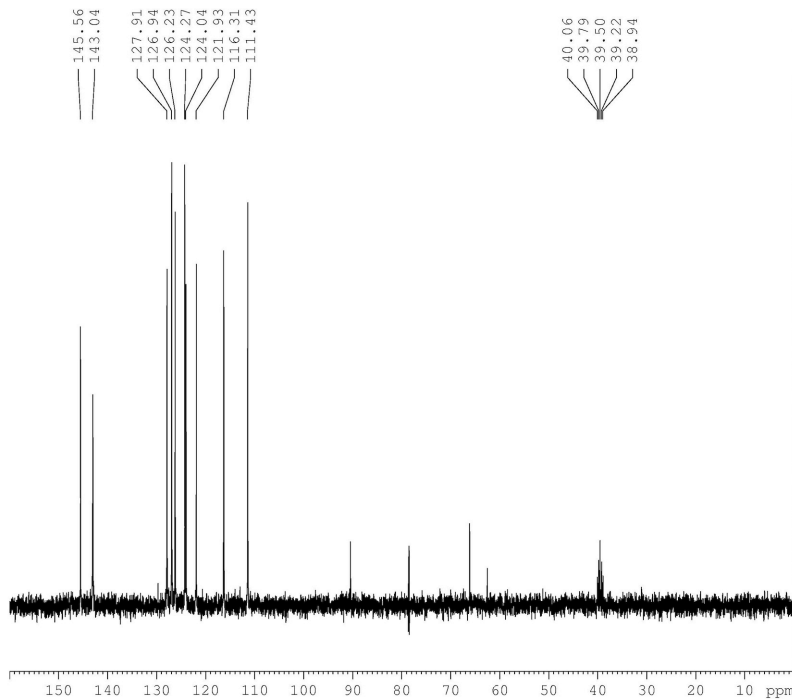
===== CHANNEL f1 =====
NUC1       13C
P1         5.80 usec
PL1        -6.00 dB
SFO1       75.4752953 MHz

===== CHANNEL f2 =====
CPDPRG2   waltz16
NUC2       1H
PCPD2     100.00 usec
PL2        -6.00 dB
PL12       17.00 dB
PL13       19.00 dB
SFO2       300.1312005 MHz

F2 - Processing parameters
SI         32768
SF         75.4677199 MHz
WDW        EM
SSB         0
LB         1.00 Hz
GB         0
PC         1.40
    
```

C13 DEPT90

C13DEPT90 DMSO D: {UIO\alexeybr} 20



```

Current Data Parameters
NAME      AB 61 exx pure
EXPNO     17
PROCNO    1

F2 - Acquisition Parameters
Date_     20100222
Time      13.57
INSTRUM   DFX500
PROBHD    5 mm QNP 1H/1
PULPROG   dept90
TD         65536
SOLVENT   DMSO
NS         1024
DS         4
SWH        17985.611 Hz
FIDRES     0.274439 Hz
AQ         1.8213508 sec
RG         16384
DW         27.800 usec
DE         6.00 usec
TE         298.2 K
CNST2     145.0000000
D1         2.0000000 sec
d2         0.0344828 sec
d12        0.0002000 sec
DELTA     0.00000738 sec
TD0        1

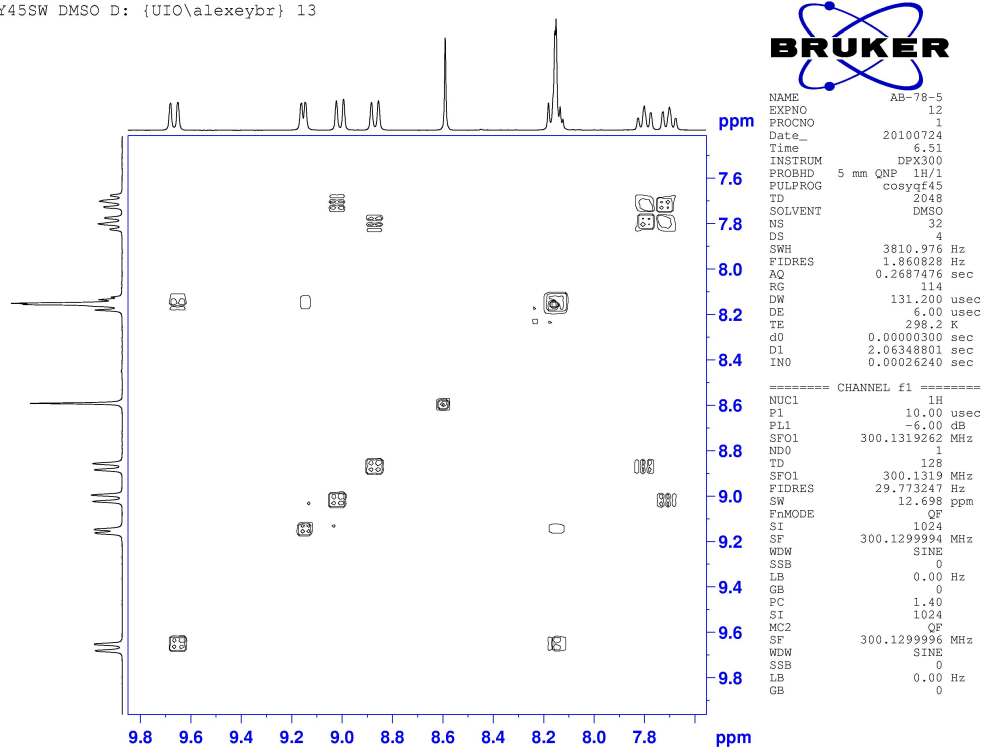
===== CHANNEL f1 =====
NUC1       13C
P1         5.80 usec
P2         11.60 usec
PL1        -6.00 dB
SFO1       75.4752953 MHz

===== CHANNEL f2 =====
CPDPRG2   waltz16
NUC2       1H
P3         10.00 usec
P4         20.00 usec
PCPD2     100.00 usec
PL2        -6.00 dB
PL12       17.00 dB
SFO2       300.1312005 MHz

F2 - Processing parameters
SI         32768
SF         75.4677199 MHz
WDW        EM
SSB         0
LB         1.00 Hz
GB         0
PC         1.40
    
```

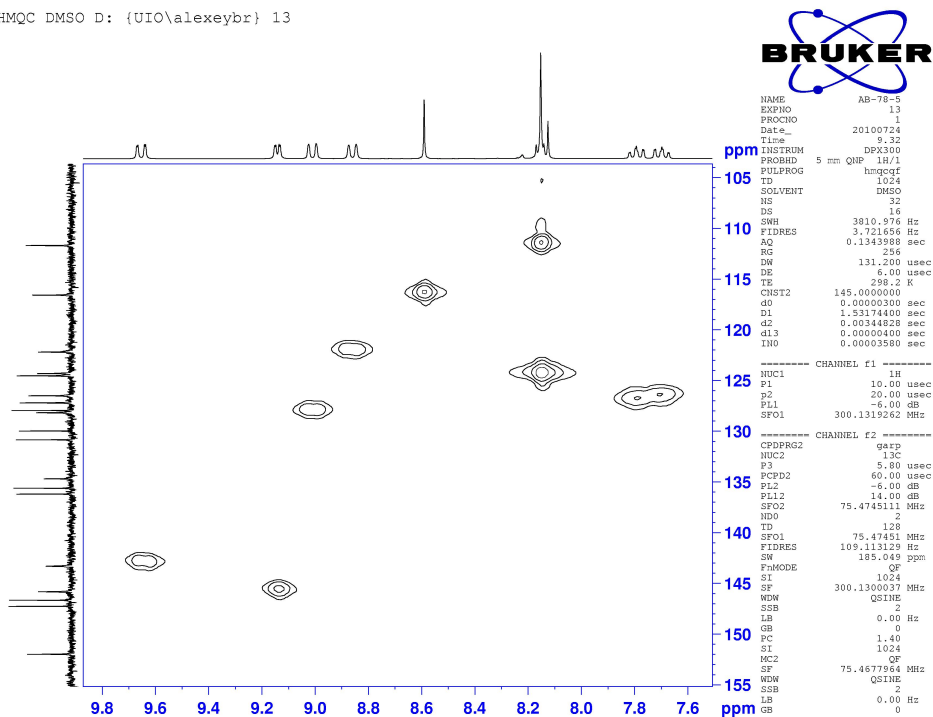
COSY45SW, expansion of the aromatic region

COSY45SW DMSO D: {UIO\alexeybr} 13



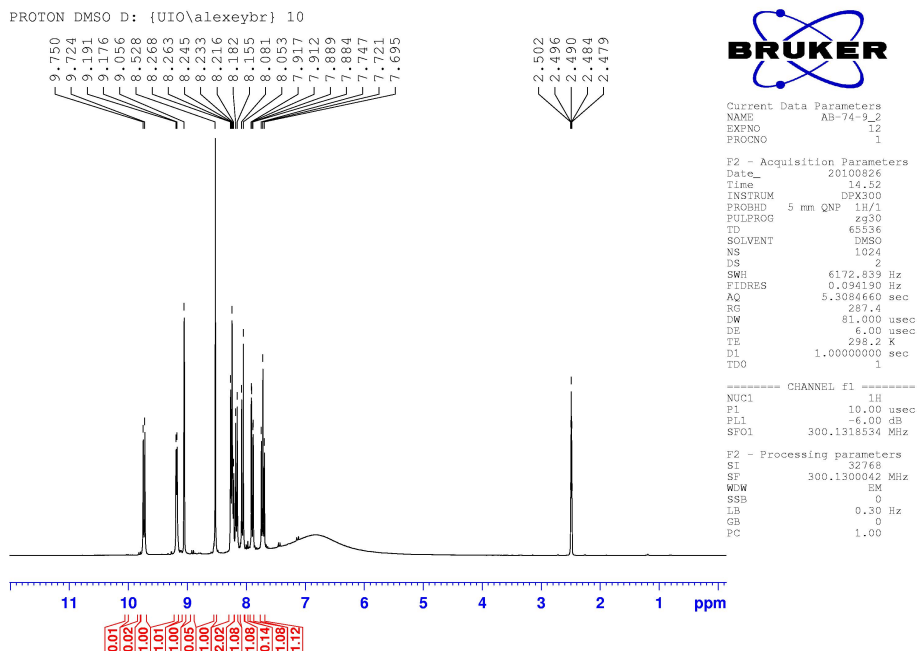
HMQC (¹H – ¹³C), expansion of the aromatic region

HMQC DMSO D: {UIO\alexeybr} 13

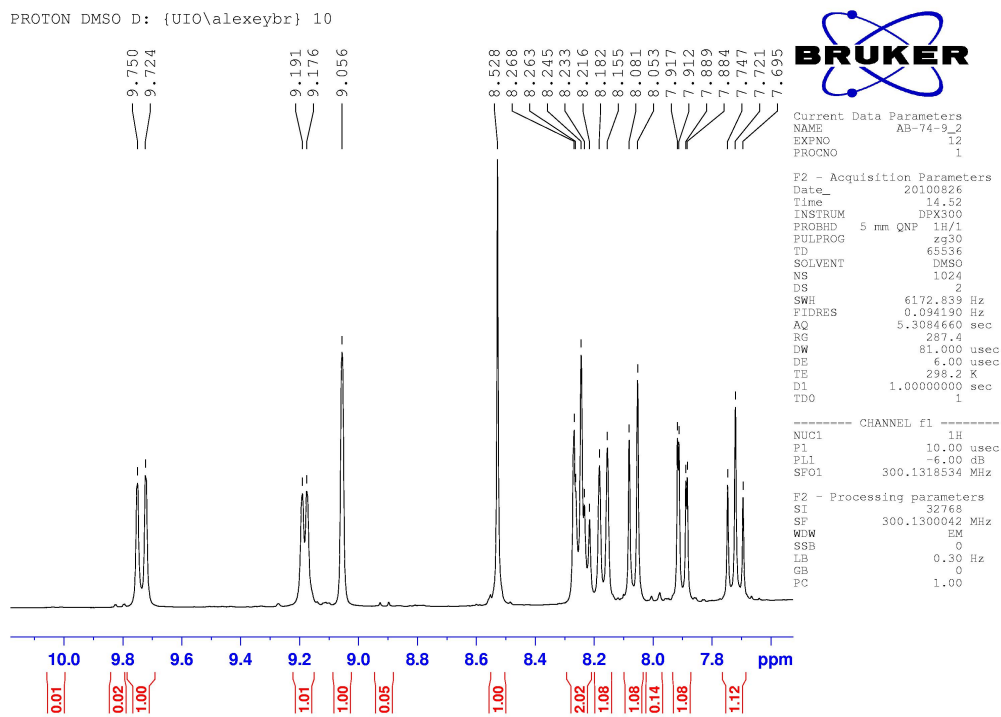


AB 74 - (E)-8-hydroxy-7-((7-sulfonaphthalen-1-yl)diazenyl)quinoline-5-sulfonic acid

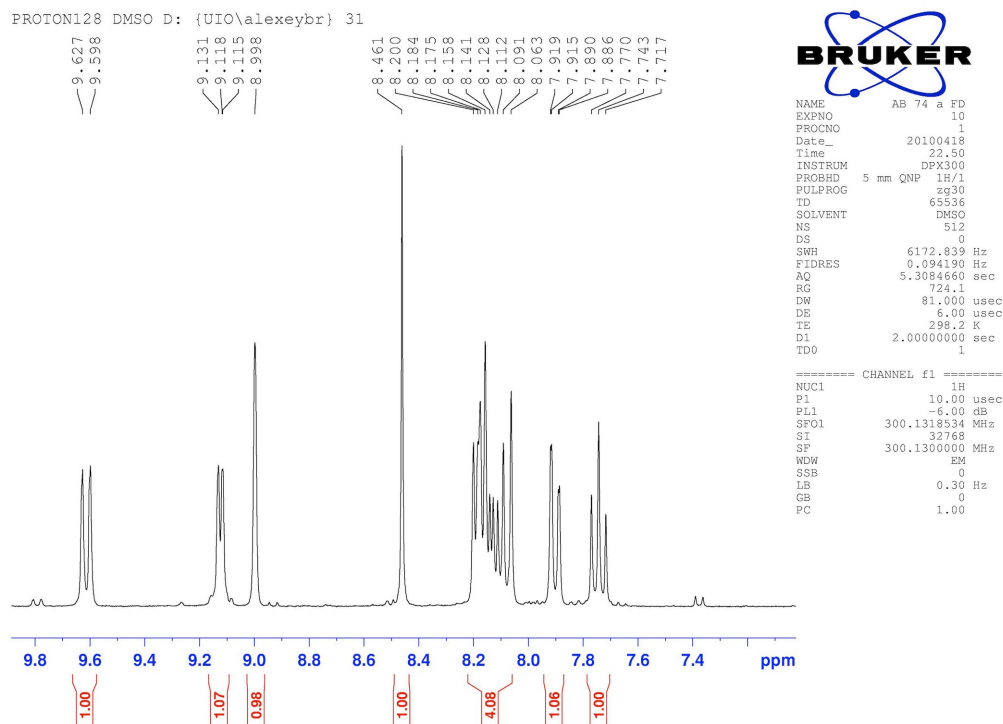
¹H NMR



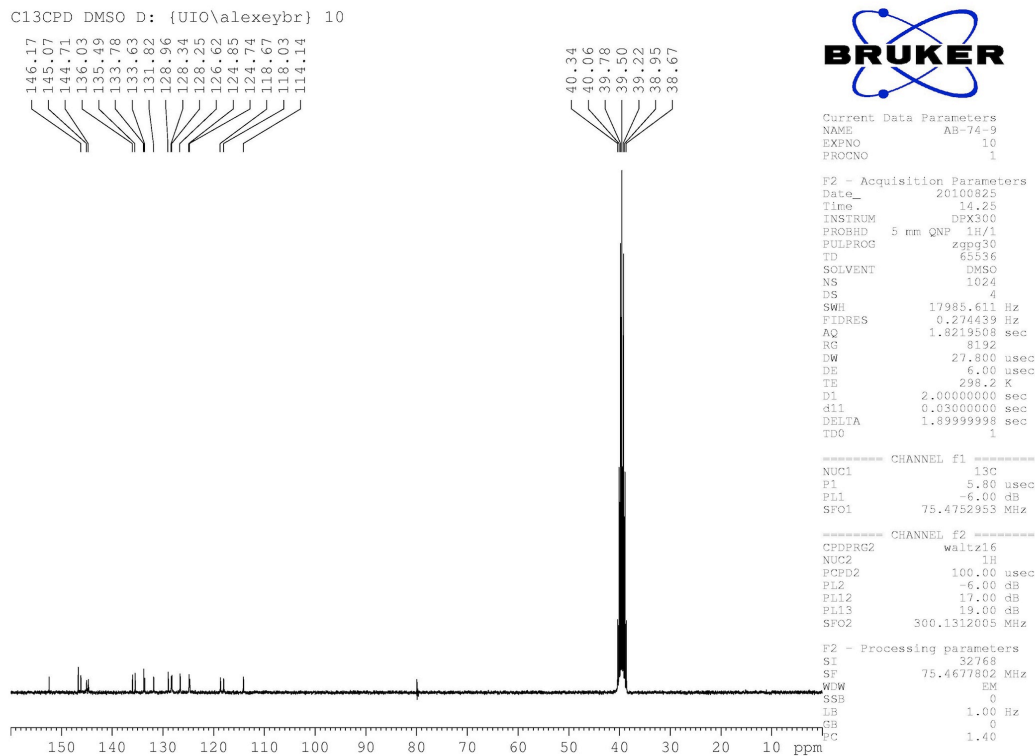
Expansion of the aromatic region



¹H NMR spectrum after freeze drying (no dioxane used in purific.), expansion of the aromatic region

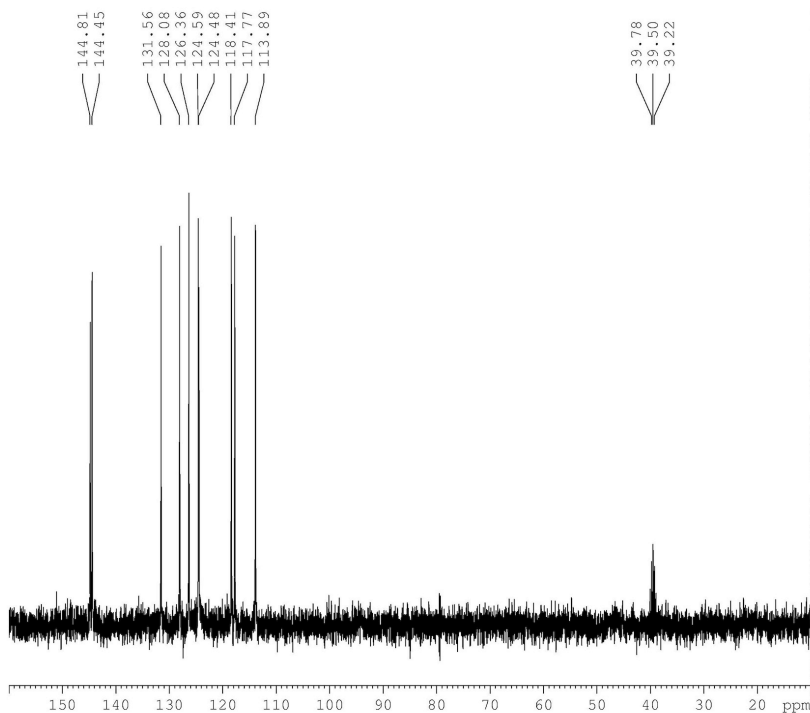


C13CPD



C13 DEPT90

C13DEPT90 DMSO D: {UIO\alexeybr} 10



Current Data Parameters
 NAME AB-74-9
 EXPNO 11
 PROCNO 1

F2 - Acquisition Parameters
 Date_ 20100825
 Time 14.59
 INSTRUM DPK300
 PROBHD 5 mm QNP 1H/1
 PULPROG dept90
 TD 65536
 SOLVENT DMSO
 NS 512
 DS 4
 SWH 17985.611 Hz
 FIDRES 0.274439 Hz
 AQ 1.8219508 sec
 RG 11585.2
 DW 27.800 usec
 DE 6.00 usec
 TE 298.2 K
 CNST2 145.000000
 D1 2.00000000 sec
 d2 0.00344828 sec
 d12 0.00000000 sec
 DELTA 0.00000738 sec
 TDD 1

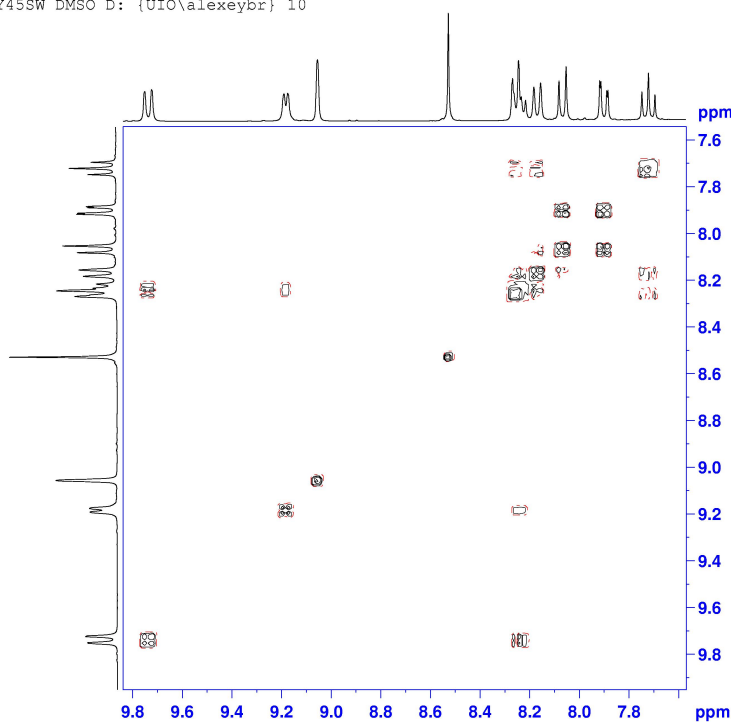
===== CHANNEL f1 =====
 NUC1 13C
 P1 5.80 usec
 p2 11.60 usec
 PL1 -6.00 dB
 SFO1 75.4752953 MHz

===== CHANNEL f2 =====
 CPDPRG2 waltz16
 NUC2 1H
 P3 10.00 usec
 p4 20.00 usec
 PCPD2 100.00 usec
 PL2 -6.00 dB
 PL12 17.00 dB
 SFO2 300.1312003 MHz

F2 - Processing parameters
 SI 32768
 SF 75.4677997 MHz
 WDW EM
 SSB 0
 LB 1.00 Hz
 GB 0
 PC 1.40

COSY45SW, expansion of the aromatic region

COSY45SW DMSO D: {UIO\alexeybr} 10

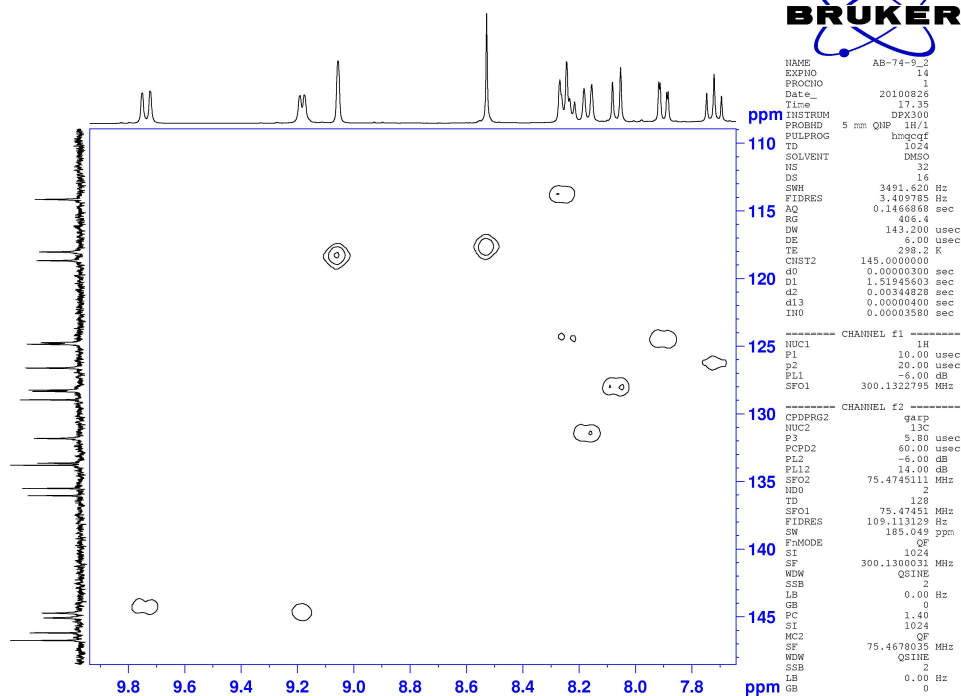


NAME AB-74-9_2
 EXPNO 13
 PROCNO 1
 Date_ 20100826
 Time 14.54
 INSTRUM DPK300
 PROBHD 5 mm QNP 1H/1
 PULPROG cosyq45
 TD 2048
 SOLVENT DMSO
 NS 32
 DS 4
 SWH 3491.620 Hz
 FIDRES 1.704893 Hz
 AQ 0.2933236 sec
 RG 2228.1
 DW 143.200 usec
 DE 6.00 usec
 TE 298.2 K
 d0 0.00000300 sec
 D1 2.03891206 sec
 INO 0.00028640 sec

===== CHANNEL f1 =====
 NUC1 1H
 P1 10.00 usec
 PL1 -6.00 dB
 SFO1 300.1322795 MHz
 NDO 1
 TD 128
 SFO1 300.1323 MHz
 FIDRES 27.278282 Hz
 SW 11.634 ppm
 PRMODE QF
 SI 1024
 SF 300.1300028 MHz
 WDW SINE
 SSB 0
 LB 0.00 Hz
 GB 0
 PC 1.40
 SI 1024
 MC2 QF
 SF 300.1300026 MHz
 WDW SINE
 SSB 0
 LB 0.00 Hz
 GB 0

HMQC (¹H – ¹³C), expansion of the aromatic region

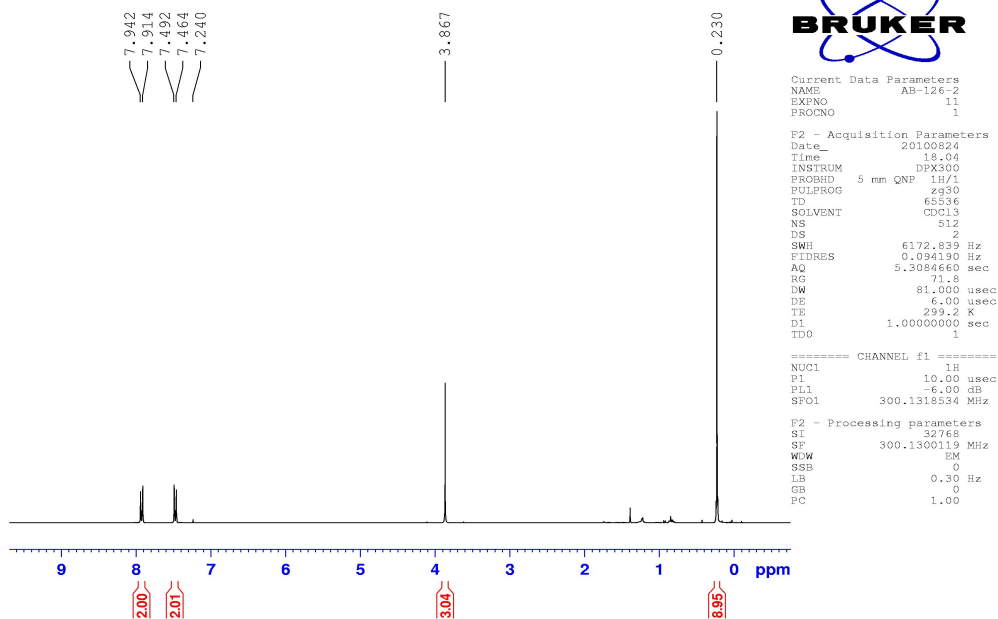
HMQC DMSO D: {UIO\alexeybr} 10



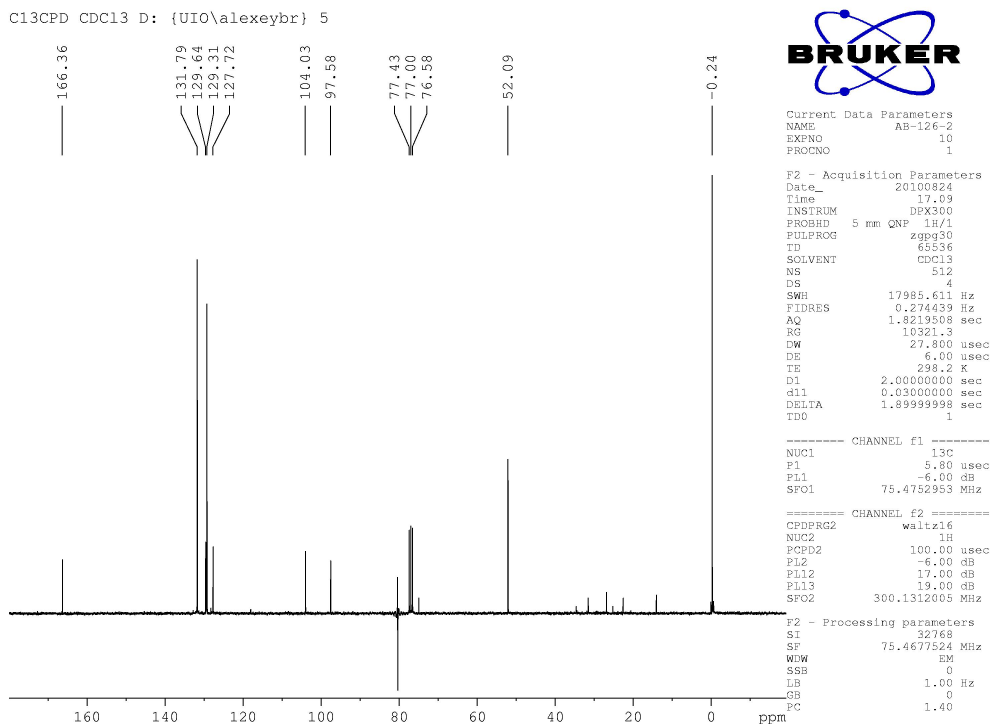
Methyl 4-((trimethylsilyl)ethynyl)benzoate

¹H NMR

PROTON CDCl3 D: {UIO\alexeybr} 5

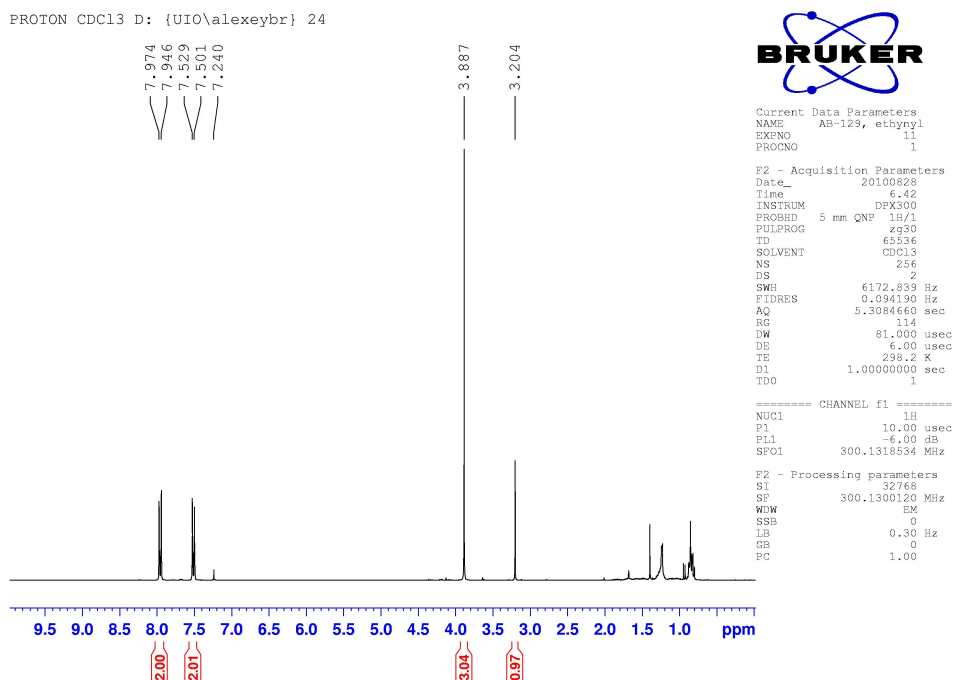


C13CPD



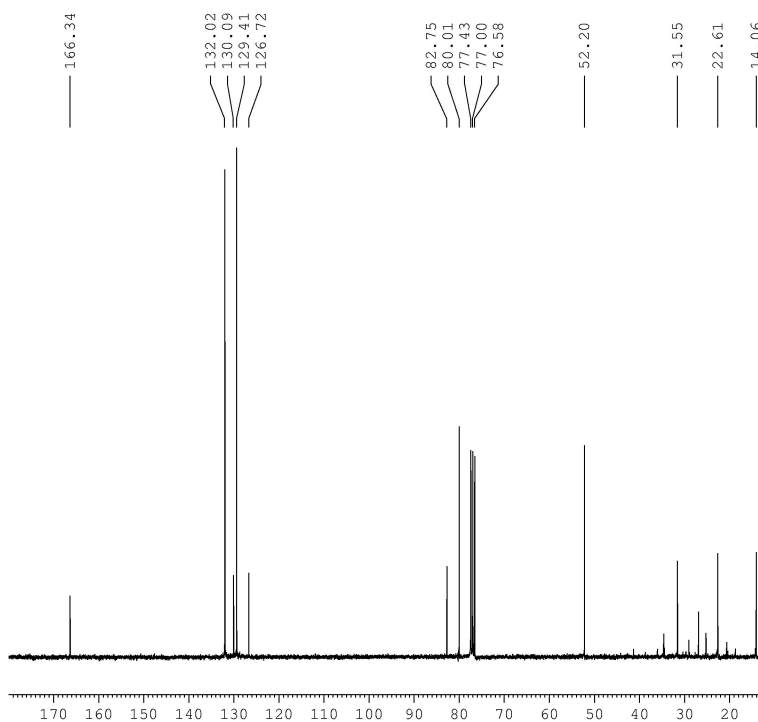
Methyl 4-ethynylbenzoate

¹H NMR



C13CPD

C13CPD CDC13 D: {UIO\alexeybr} 24



```

Current Data Parameters
NAME      AB-129, ethnyyl
EXPNO    10
PROCNO   1

F2 - Acquisition Parameters
Date_    20100828
Time     6.15
INSTRUM  DFX300
PROBHD   5 mm QNP 1H/1
PULPROG  zgpg30
TD       65536
SOLVENT  CDCl3
NS       512
DS       4
SWH      17985.611 Hz
FIDRES   0.274439 Hz
AQ       1.8219508 sec
RG       11585.2
DW       27.800 usec
DE       6.00 usec
TE       298.2 K
D1       2.0000000 sec
d11      0.0300000 sec
DELTA    1.8999999 sec
TD0      1

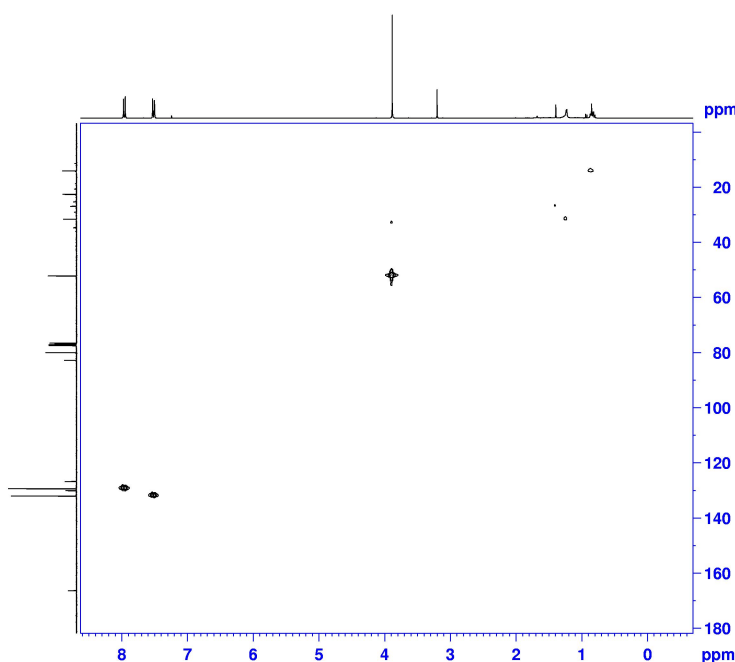
----- CHANNEL f1 -----
NUC1     13C
P1       5.80 usec
PL1      -6.00 dB
SFO1     75.4752953 MHz

----- CHANNEL f2 -----
CPDPRG2  waltz16
NUC2     1H
PCPD2   100.00 usec
PL2      -6.00 dB
PL12     17.00 dB
PL13     19.00 dB
SFO2     300.1312005 MHz

F2 - Processing parameters
SI       32768
SF       75.4677516 MHz
WDW      EM
SSB      0
LB       1.00 Hz
GB       0
PC       1.40
    
```

HMQC (¹H - ¹³C)

HMQC CDC13 D: {UIO\alexeybr} 24



```

Current Data Parameters
NAME      AB-129, ethnyyl
EXPNO    13
PROCNO   1

F2 - Acquisition Parameters
Date_    20100828
Time     7.45
INSTRUM  DFX300
PROBHD   5 mm QNP 1H/1
PULPROG  hmcqcf
TD       1024
SOLVENT  CDCl3
NS       32
DS       16
SWH      2794.421 Hz
FIDRES   2.730880 Hz
AQ       0.1931412 sec
RG       181
DW       178.800 usec
DE       6.00 usec
TE       298.2 K
CST2    145.000000 sec
d0       0.0000000 sec
D1       1.4825919 sec
d2       0.00346828 sec
d13      0.0000000 sec
IN0      0.0000350 sec

----- CHANNEL f1 -----
NUC1     1H
P1       10.00 usec
PL1      -6.00 dB
SFO1     300.1312005 MHz

----- CHANNEL f2 -----
CPDPRG2  gddp
NUC2     13C
PCPD2   60.00 usec
PL2      -6.00 dB
PL12     19.00 dB
SFO2     75.4751511 MHz

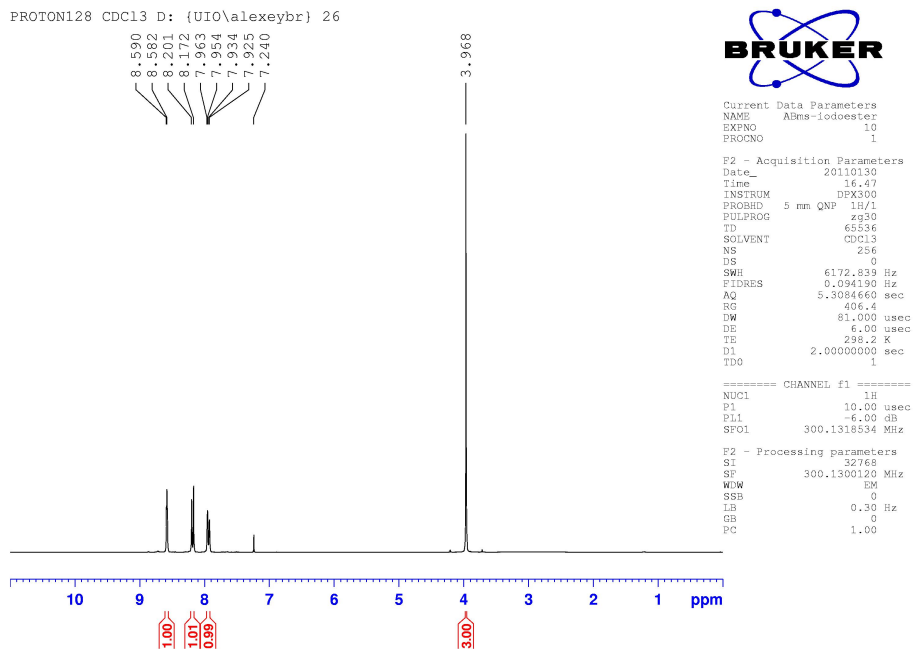
F1 - Acquisition parameters
ND0      2
TD       128
SFO1     75.47451 MHz
FIDRES   108.113129 Hz
SW       185.000 ppm
FWD0DE   0

F2 - Processing parameters
SI       1024
SF       300.130117 MHz
WDW      SINE
SSB      2
LB       0.00 Hz
GB       0
PC       1.40

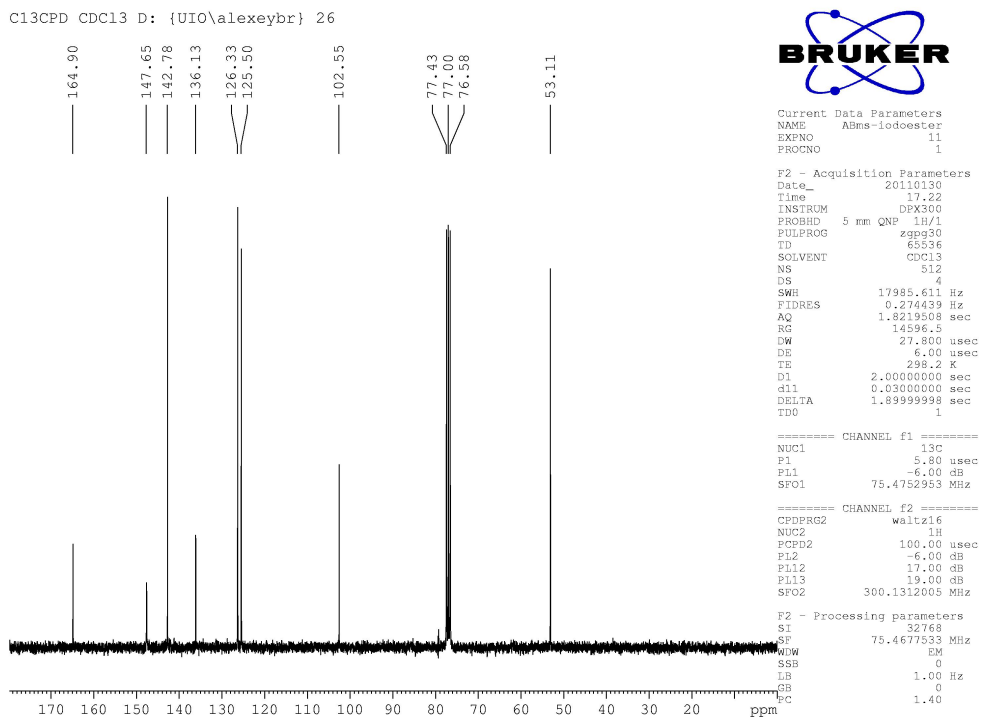
F1 - Processing parameters
SI       1024
SF       75.4677709 MHz
WDW      SINE
SSB      0
LB       0.00 Hz
GB       0
    
```


Methyl 2-iodo-5-nitrobenzoate

¹H NMR

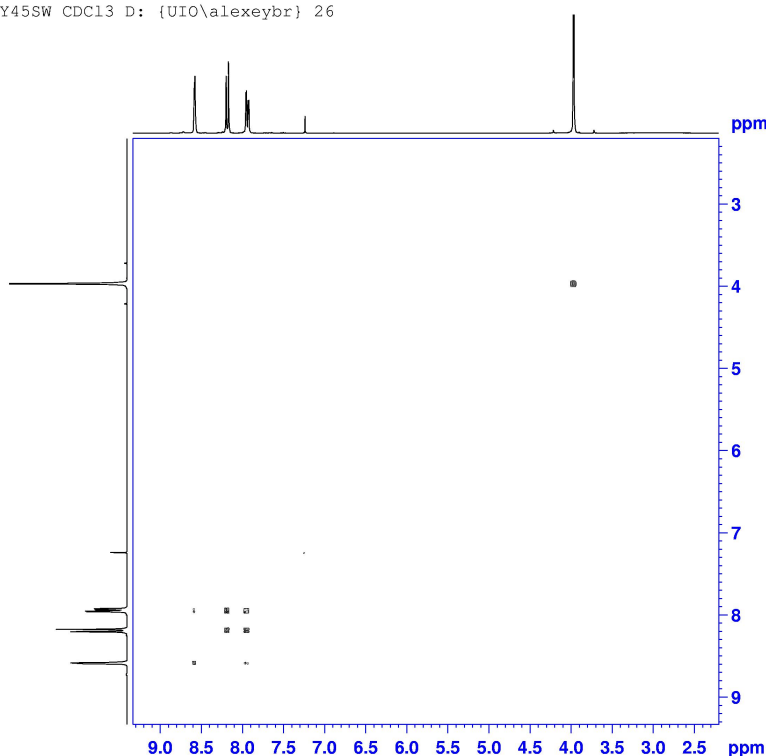


¹³C NMR



COSY45SW

COSY45SW CDC13 D: {UIO\alexeybr} 26



```

Current Data Parameters
NAME      ABms-Iodoester
EXENO     14
PROCNO    1

F2 - Acquisition Parameters
Date_     20110130
Time      17.34
INSTRUM   DFX300
PROBHD    5 mm QNP 1H/1
PULPROG   cosyg45
TD         2048
SOLVENT   CDC13
NS         8
DS         4
SWH        2140.411 Hz
FIDRES     1.045123 Hz
AQ         0.4784628 sec
RG         322.5
DW         233.600 usec
DE         6.00 usec
TE         298.2 K
d0         0.00000300 sec
d1         1.82254300 sec
INO        0.00046720 sec

===== CHANNEL f1 =====
NUC1      1H
P1        10.00 usec
PL1       -6.00 dB
SFO1      300.1317416 MHz

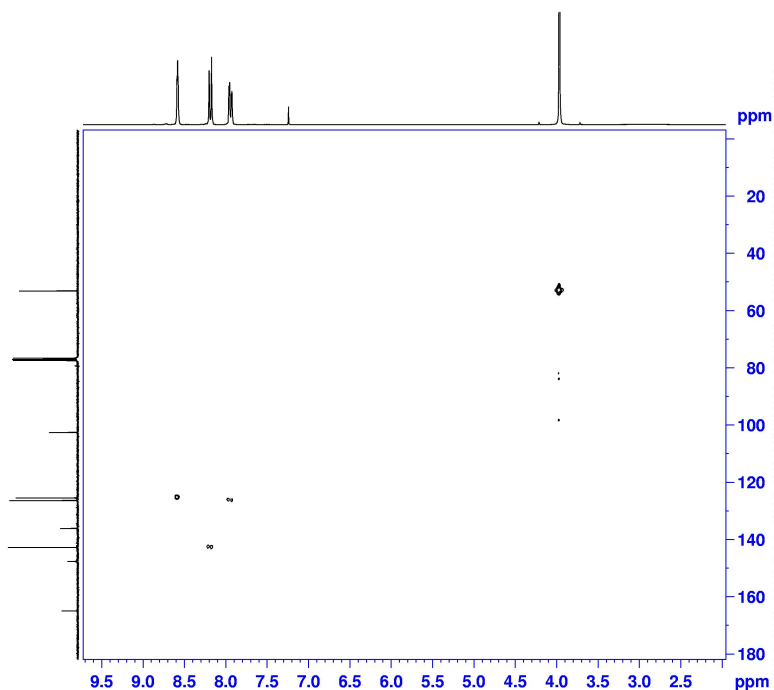
F1 - Acquisition parameters
ND0       1
TD        128
SFO1      300.1317 MHz
FIDRES    16.721960 Hz
SW        7.132 ppm
FAMODE    QF

F2 - Processing parameters
SI        1024
SF        300.1300109 MHz
WDW       SINE
SSB       0
LB        0.00 Hz
GB        0
PC        1.40

F1 - Processing parameters
SI        1024
MC2       QF
SF        300.1300109 MHz
WDW       SINE
SSB       0
LB        0.00 Hz
GB        0
  
```

HMQC (¹H - ¹³C)

HMQC CDC13 D: {UIO\alexeybr} 26



```

Current Data Parameters
NAME      ABms-Iodoester
EXENO     15
PROCNO    1

F2 - Acquisition Parameters
Date_     20110130
Time      18.16
INSTRUM   DFX300
PROBHD    5 mm QNP 1H/1
PULPROG   hmqc2f
TD         1024
SOLVENT   CDC13
NS         16
DS         16
SWH        2332.050 Hz
FIDRES     2.277451 Hz
AQ         0.2195950 sec
RG         724.1
DW         214.400 usec
DE         6.00 usec
TE         298.2 K
d0         0.00000300 sec
d1         1.44872741 sec
d2         0.00344828 sec
d13       0.00000400 sec
INO        0.00003880 sec

===== CHANNEL f1 =====
NUC1      1H
P1        10.00 usec
PL1       -6.00 dB
SFO1      300.1317648 MHz

===== CHANNEL f2 =====
CPDPRG2   srep
NUC2      13C
P2        5.80 usec
PL2       -6.00 dB
SFO2      75.4745111 MHz

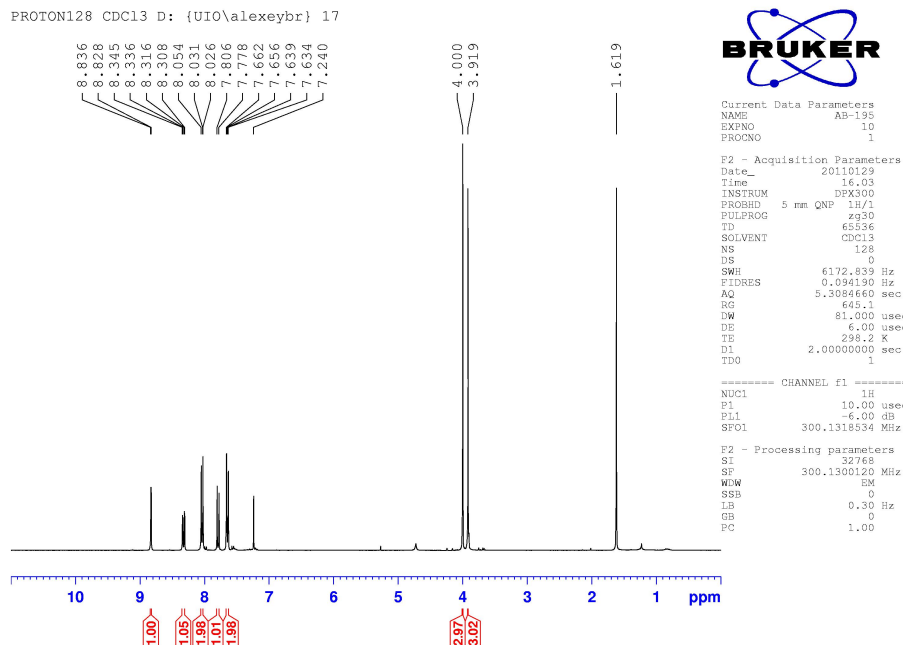
F1 - Acquisition parameters
ND0       2
TD        128
SFO1      75.47451 MHz
FIDRES    109.113129 Hz
SW        185.049 ppm
FAMODE    QF

F2 - Processing parameters
SI        1024
SF        300.1300119 MHz
WDW       SINE
SSB       2
LB        0.00 Hz
GB        0
PC        1.40

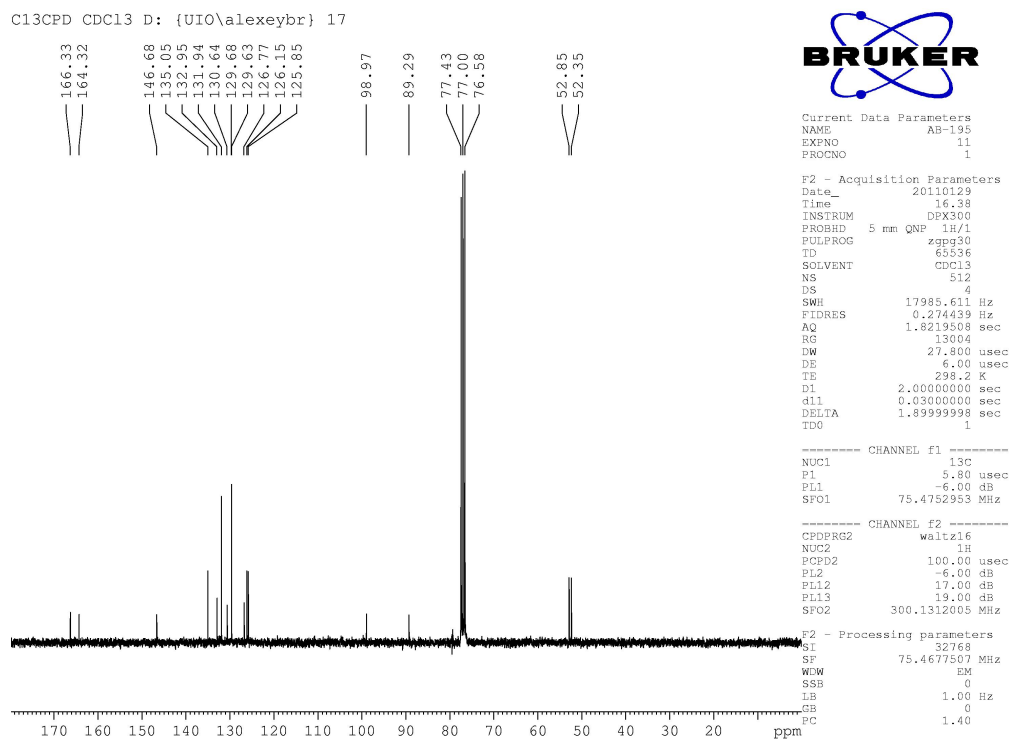
F1 - Processing parameters
SI        1024
MC2       QF
SF        75.4677636 MHz
WDW       QFINT
SSB       0
LB        0.00 Hz
GB        0
  
```

Methyl 2-((4-(methoxycarbonyl)phenyl)ethynyl)-5-nitrobenzoate

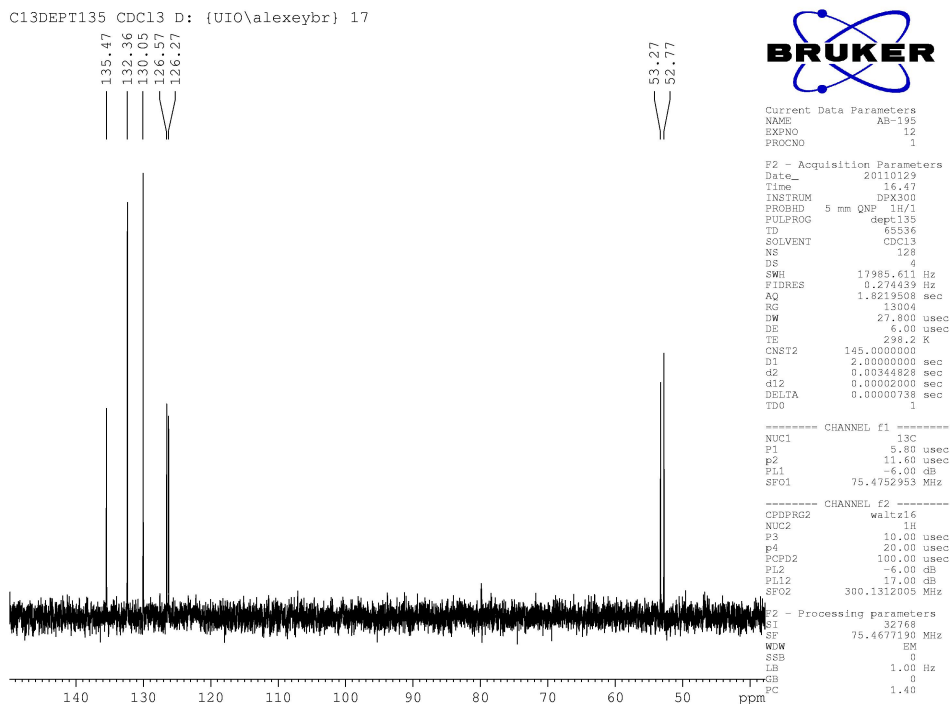
¹H NMR



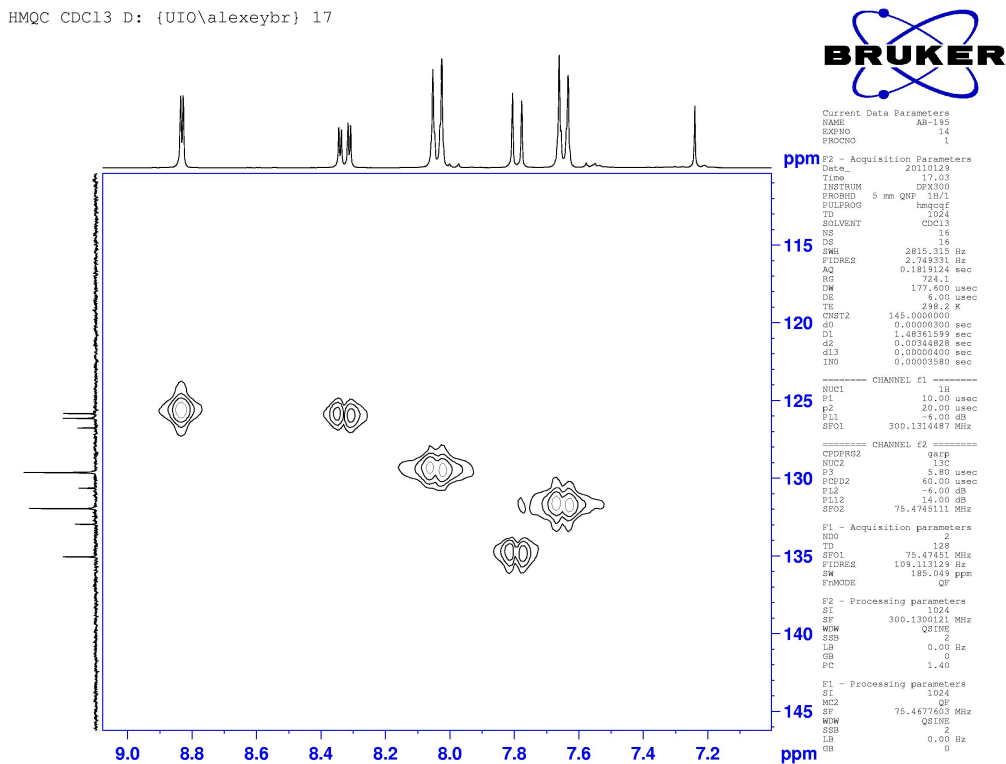
C13CPD



C13 DEPT135

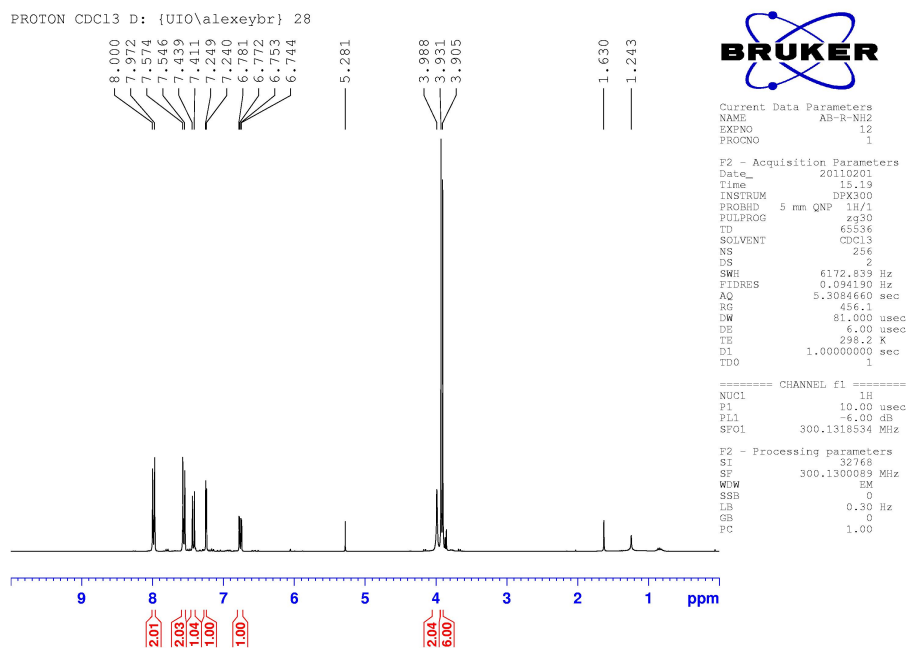


HMQC (¹H - ¹³C), expansion of the aromatic region

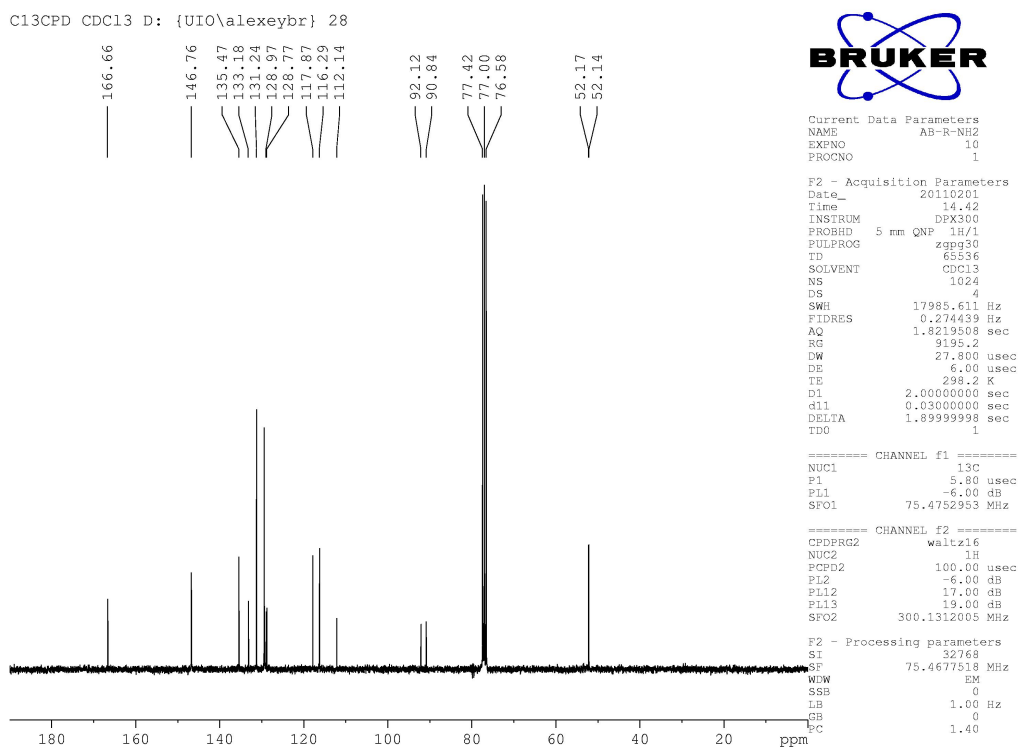


Methyl 5-amino-2-((4-(methoxycarbonyl)phenyl)ethynyl)benzoate

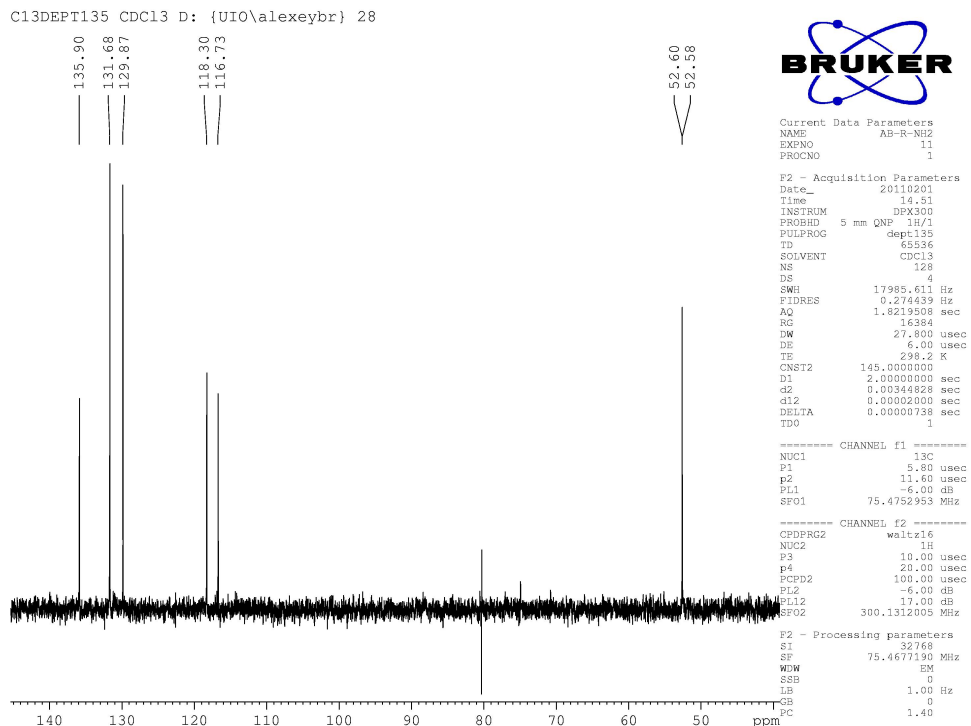
¹H NMR



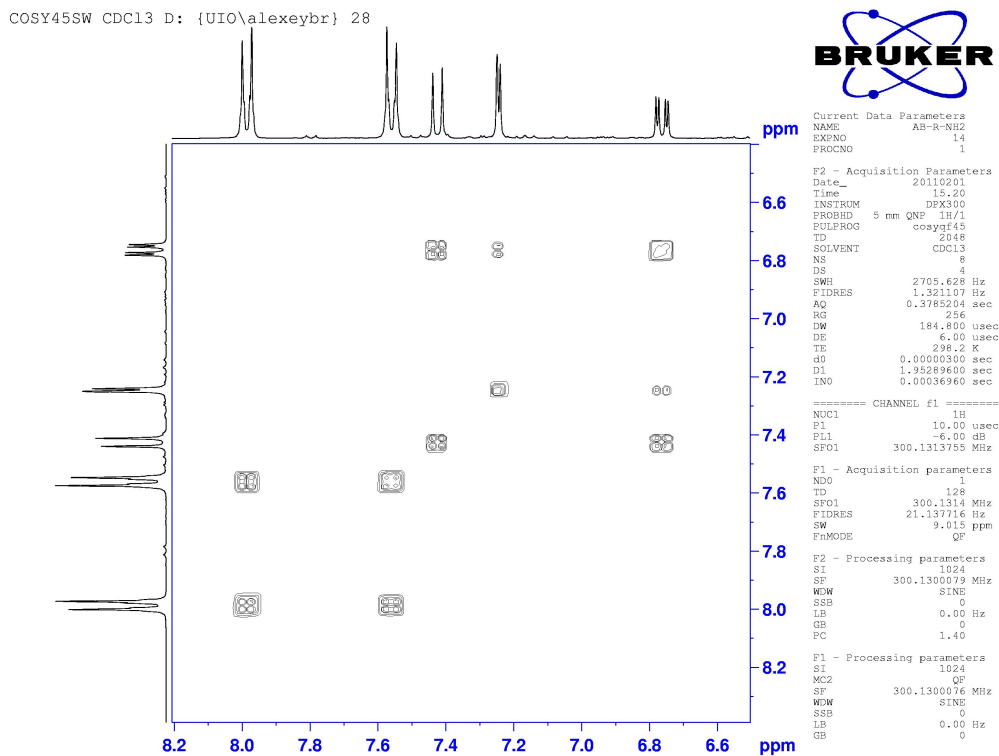
¹³C NMR



C13 DEPT135

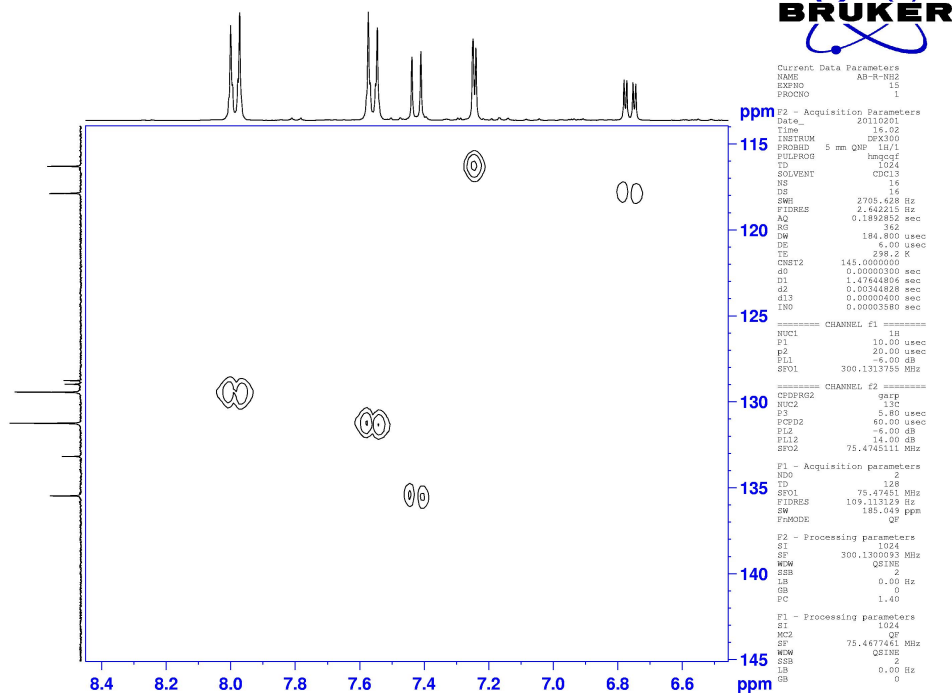


COSY45SW, expansion of the aromatic region



HMQC ($^1\text{H} - ^{13}\text{C}$), expansion of the aromatic region

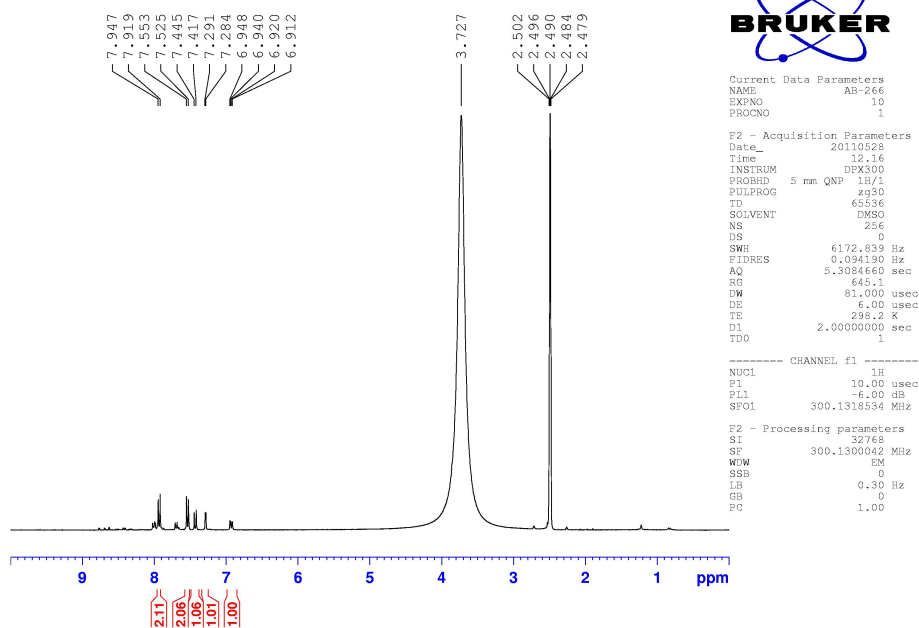
HMQC CDCl3 D: {UIO\alexeybr} 28



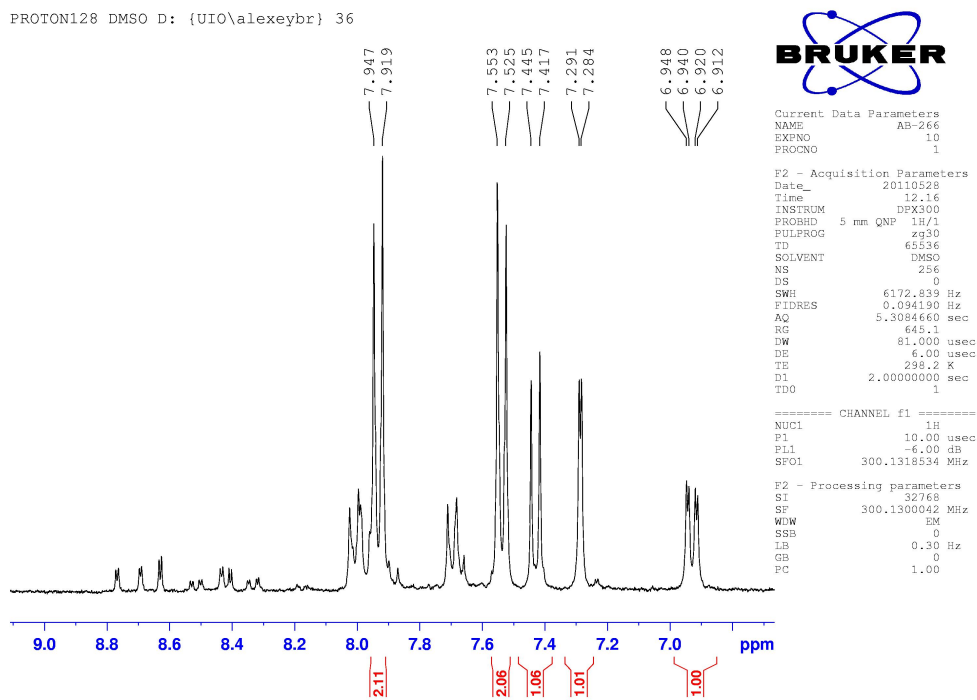
5-Amino-2-((4-carboxyphenyl)ethynyl)benzoic acid

^1H NMR

PROTON128 DMSO D: {UIO\alexeybr} 36

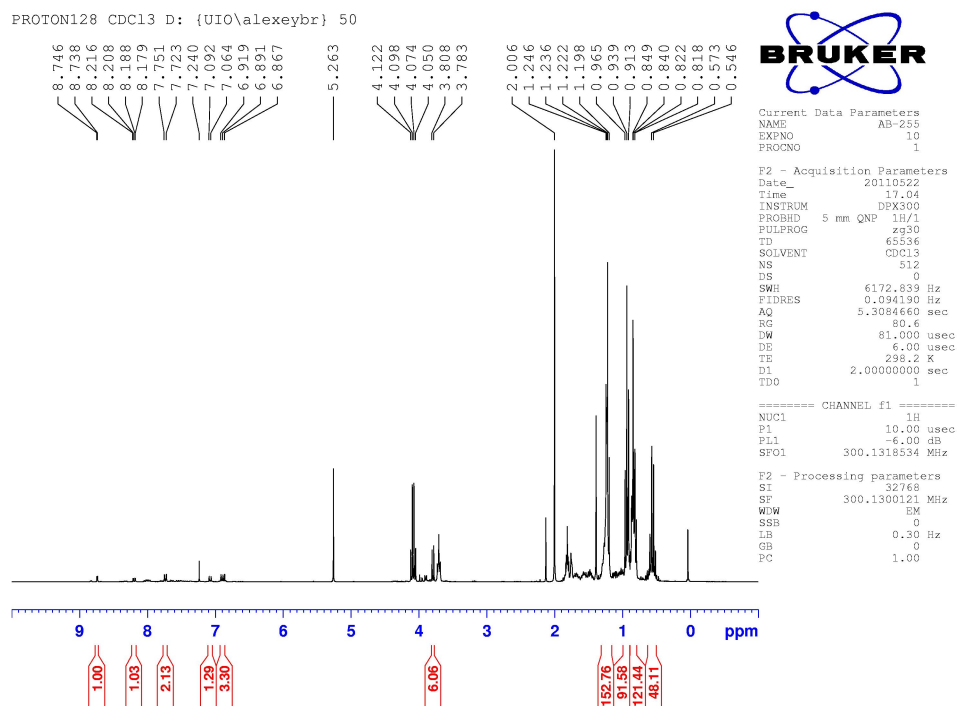


Expansion of the aromatic region

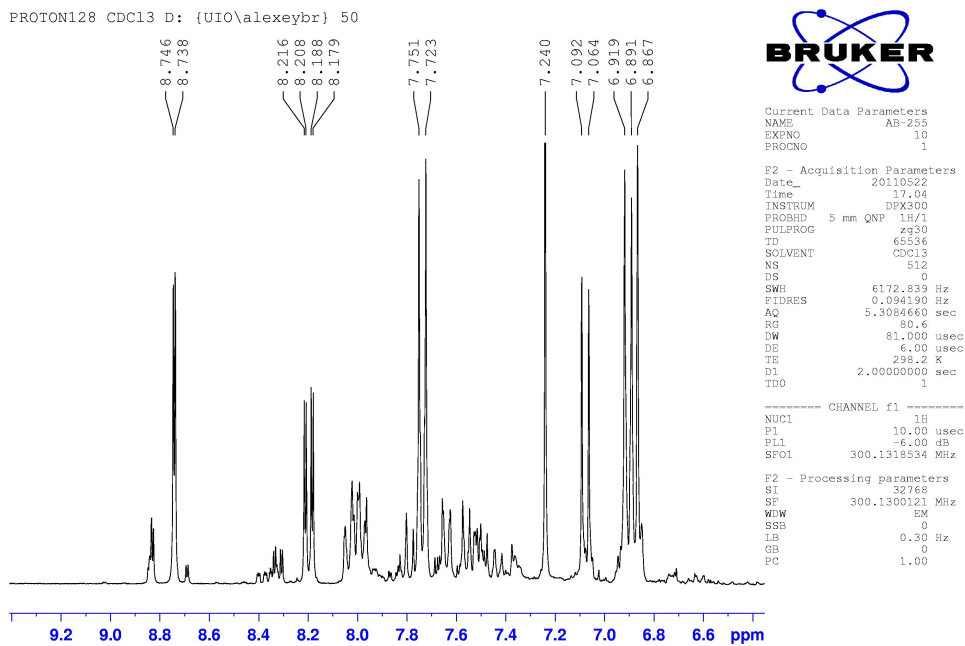


(E)-Methyl 2-(2-(4-(methoxycarbonyl)phenyl)-2-(triethylsilyl)vinyl)-5-nitrobenzoate

¹H NMR

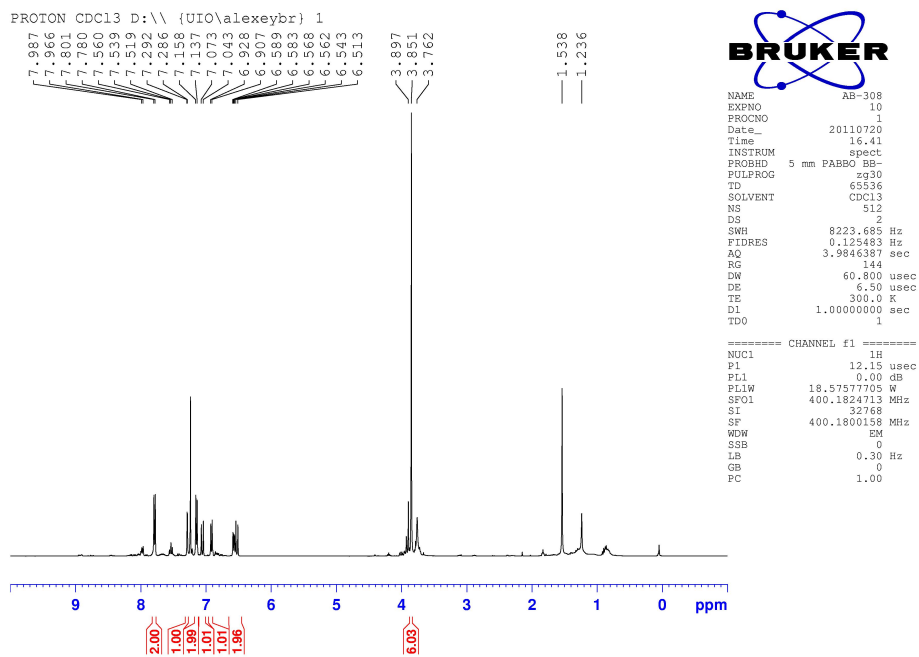


Expansion of the aromatic region

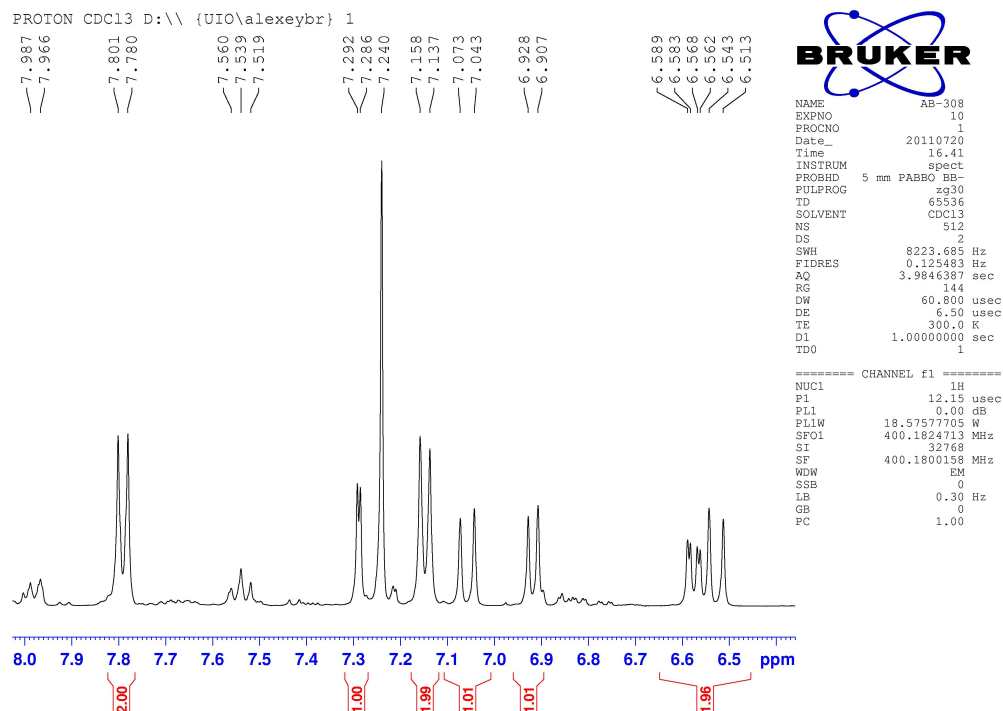


(Z)-Methyl 5-amino-2-(4-(methoxycarbonyl)styryl)benzoate

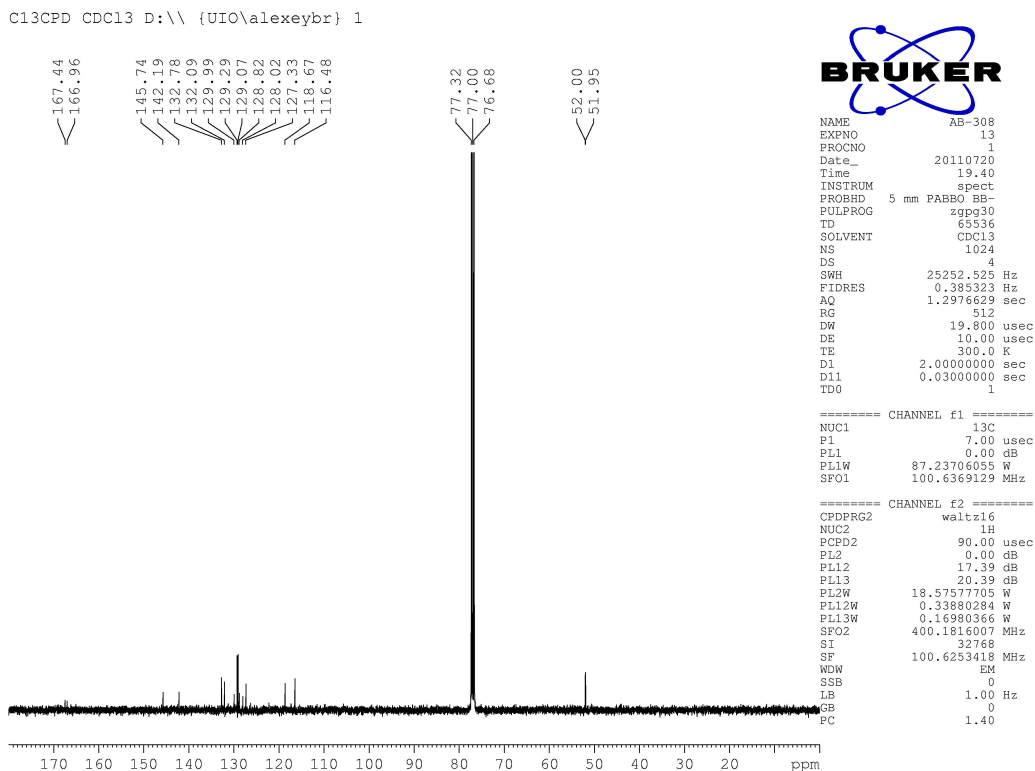
¹H NMR



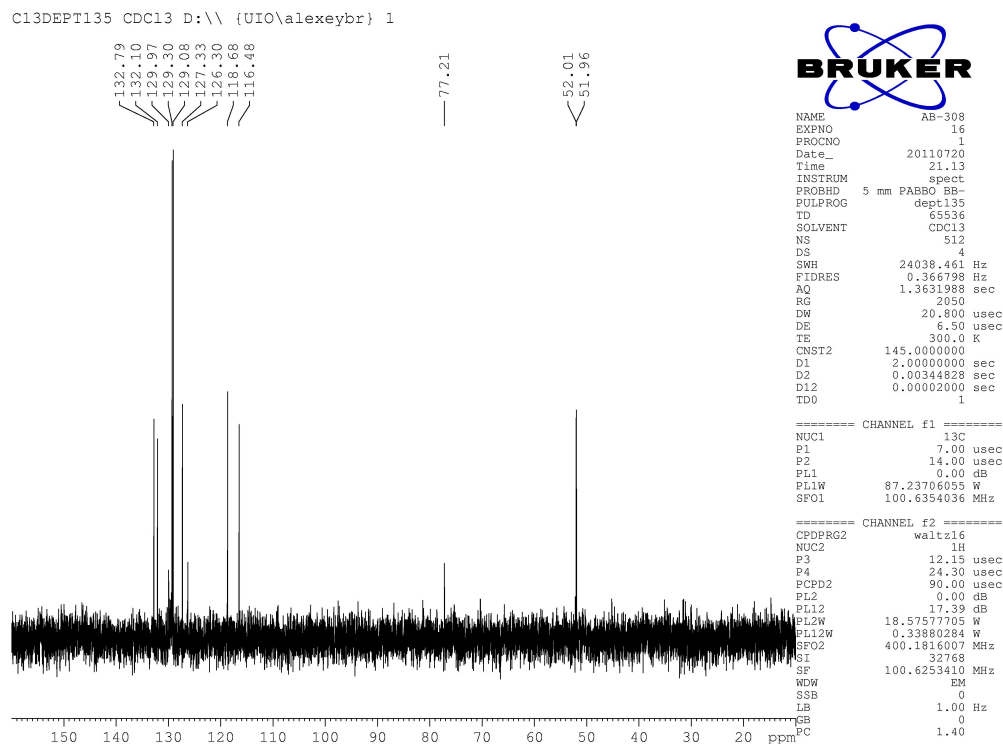
Expansion of the aromatic region



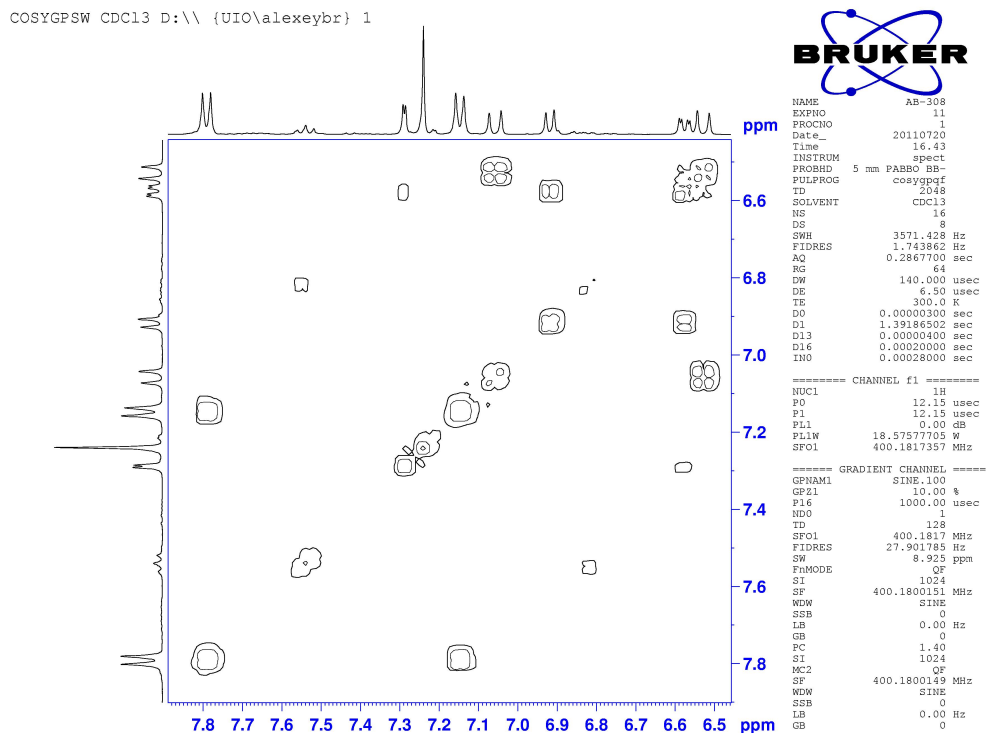
C13CPD



C13 DEPT135

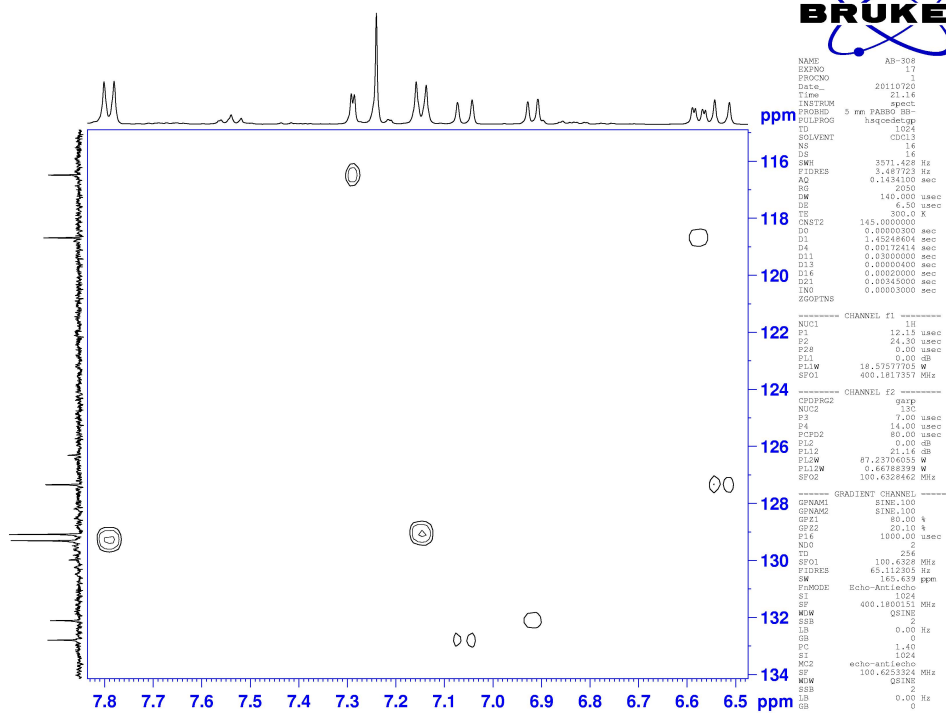


COSY45SW, expansion of the aromatic region



HMQC ($^1\text{H} - ^{13}\text{C}$), expansion of the aromatic region

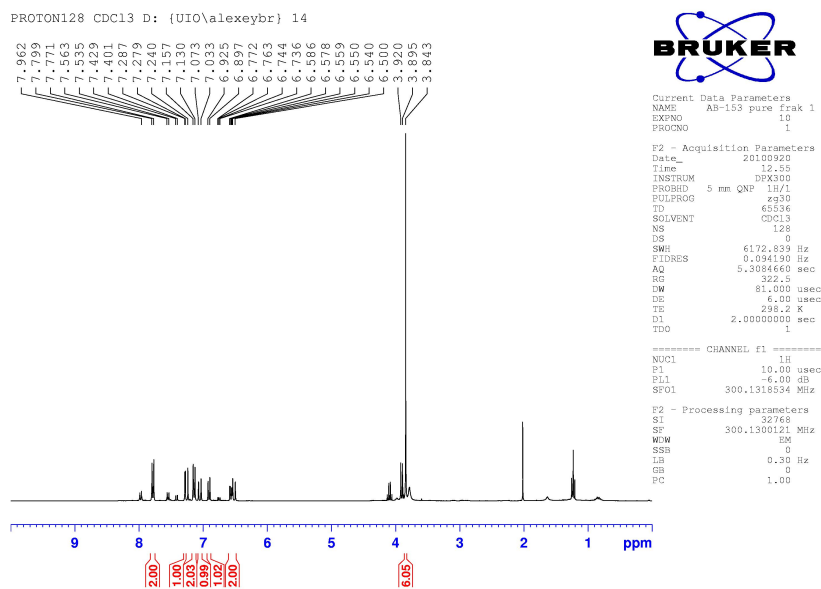
HSQCEDTGP CDC13 D:\{UIO\alexeybr} 1



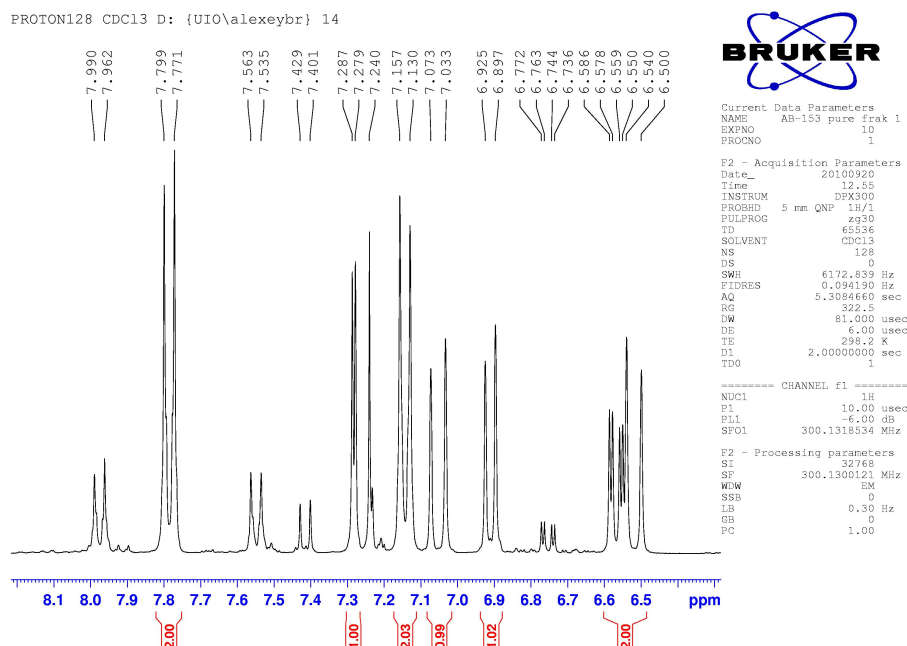
Attachment 2

Miscellaneous NMR spectra

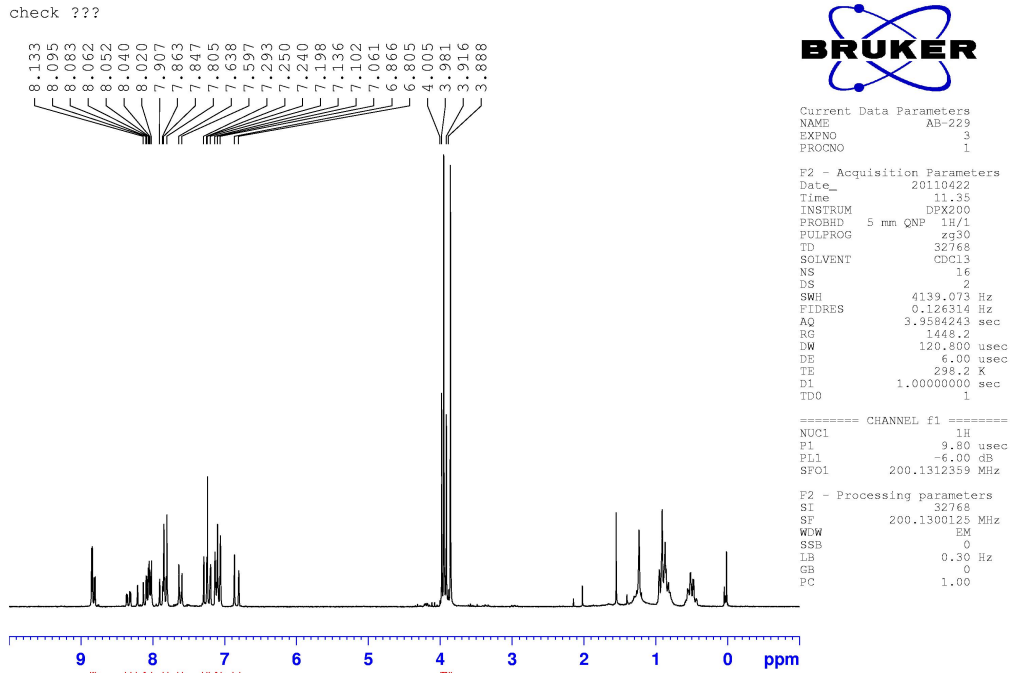
Spectrum 1



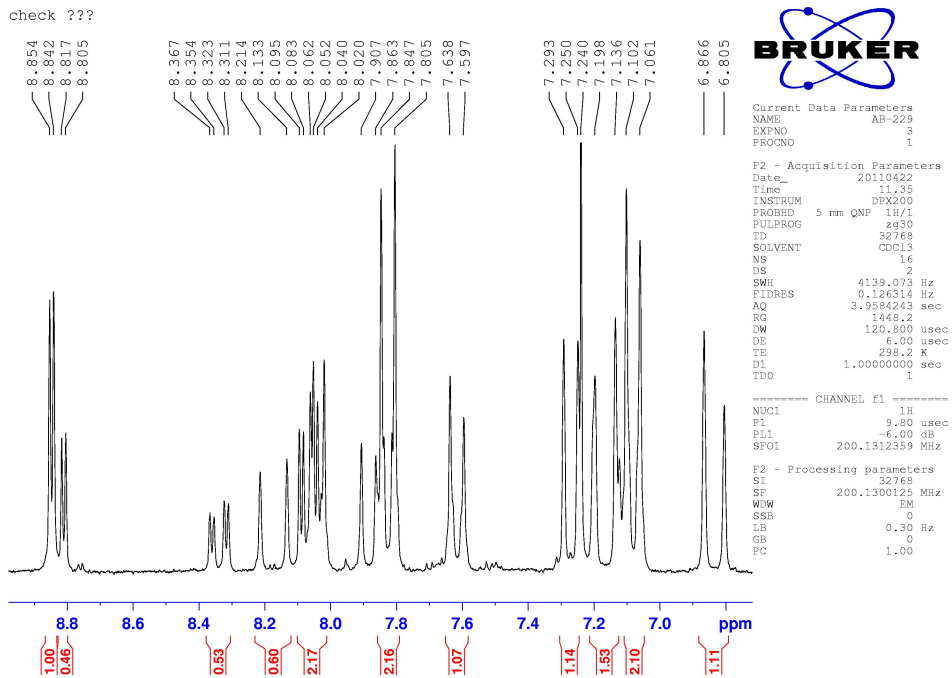
Spectrum 2, expansion of the aromatic region from spectrum 1



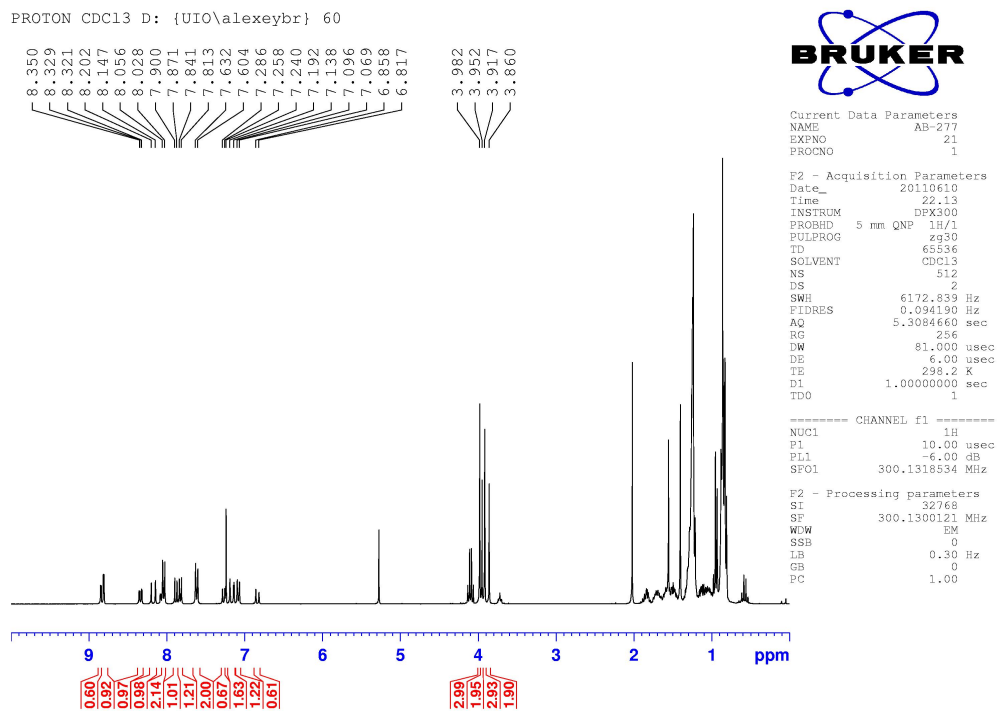
Spectrum 3



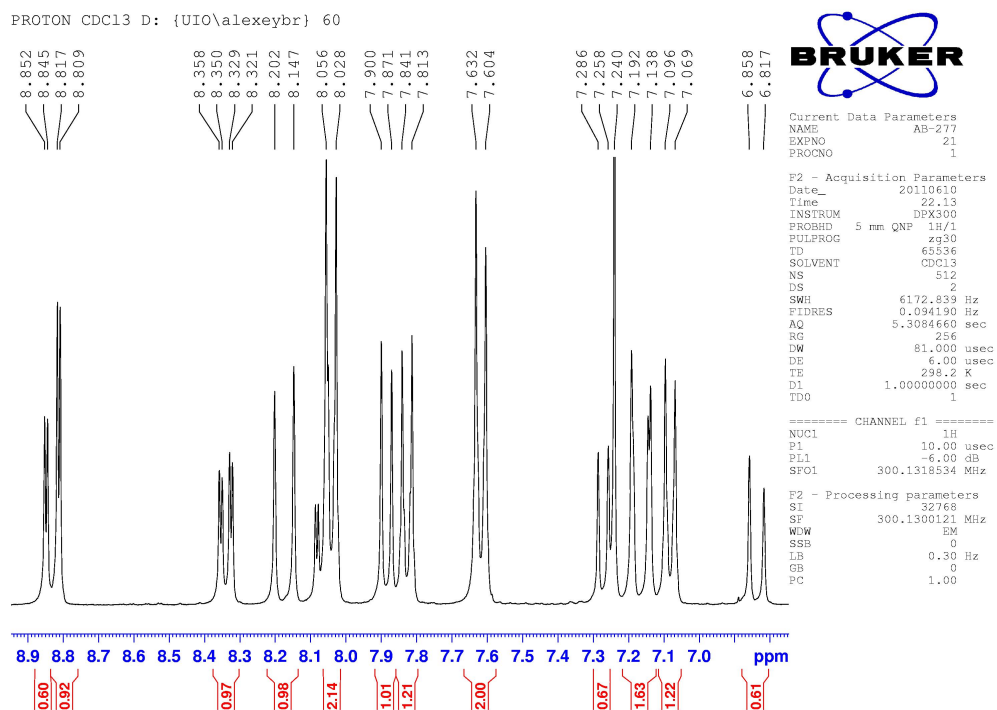
Spectrum 4, expansion of the aromatic region from spectrum 3



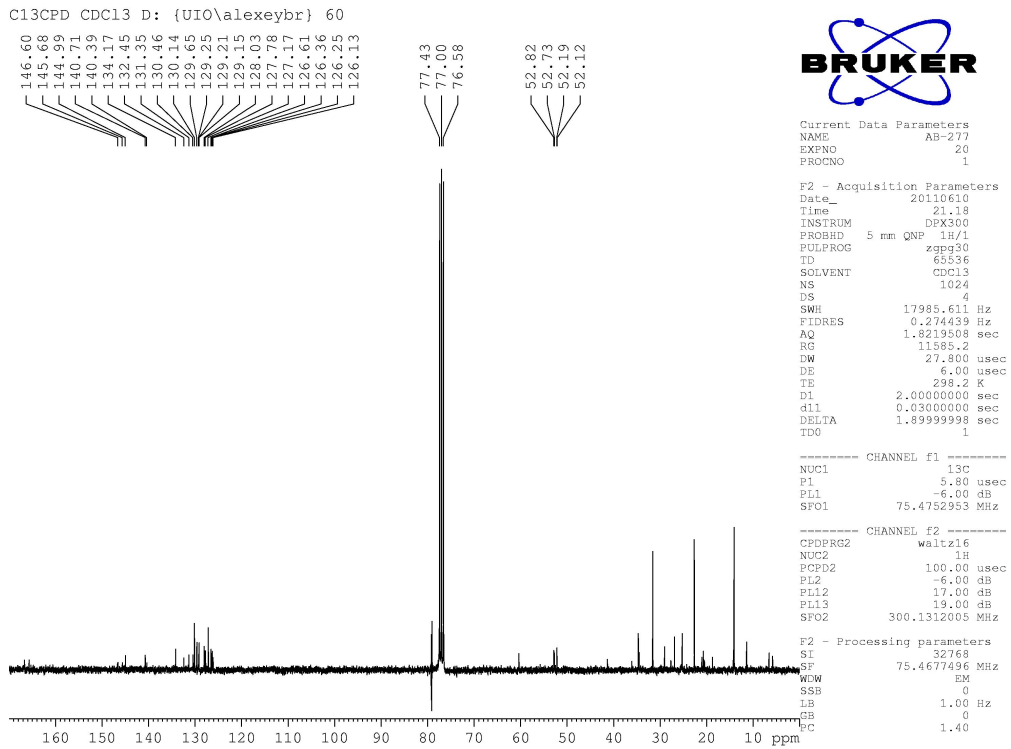
Spectrum 5



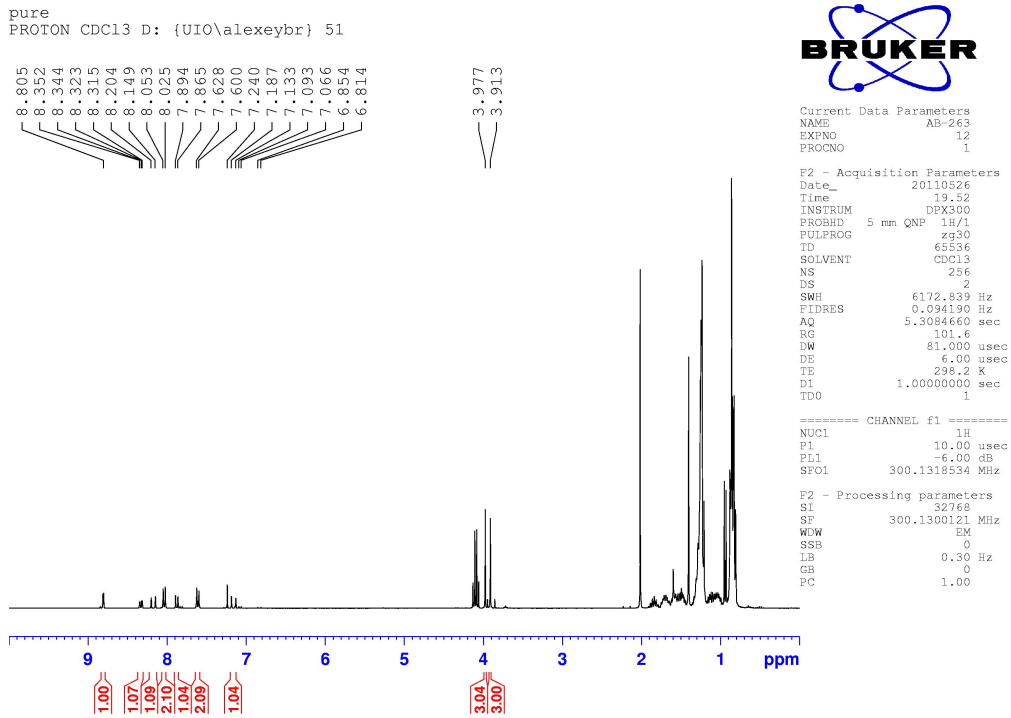
Spectrum 6, expansion of the aromatic region from spectrum 5



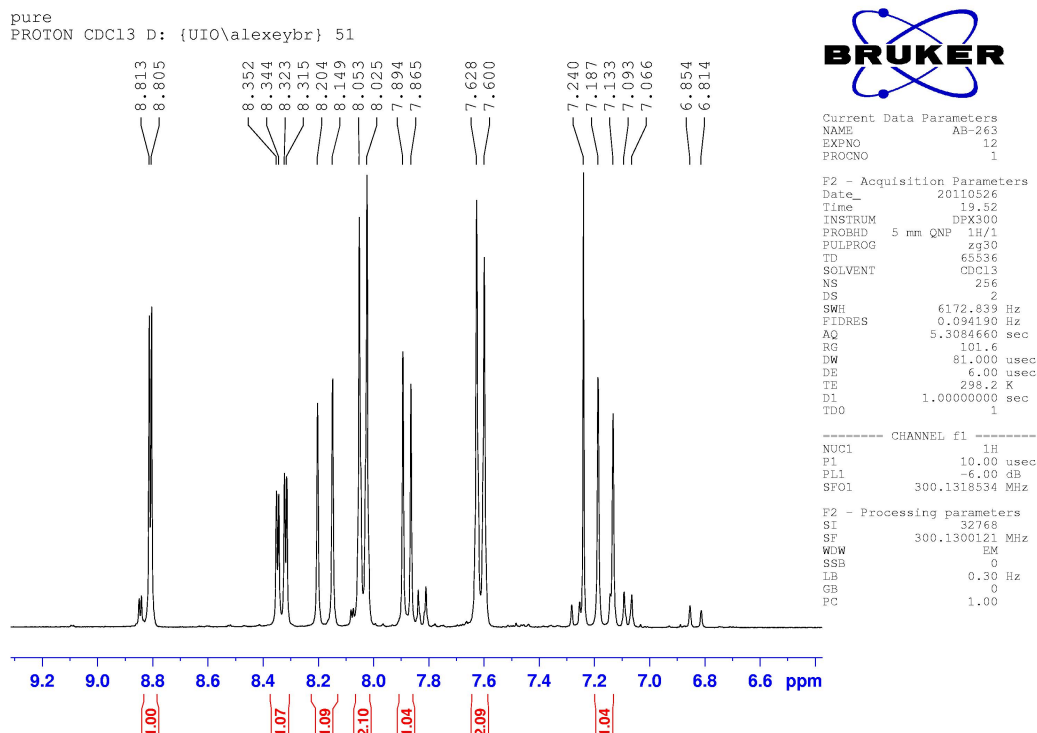
Spectrum 7



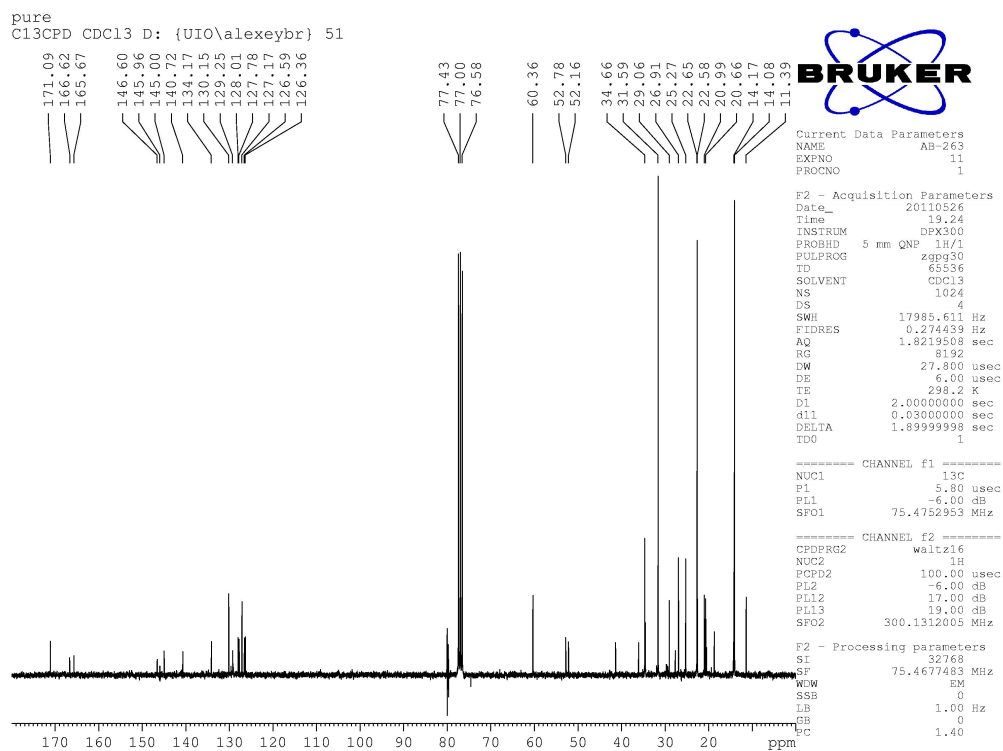
Spectrum 8



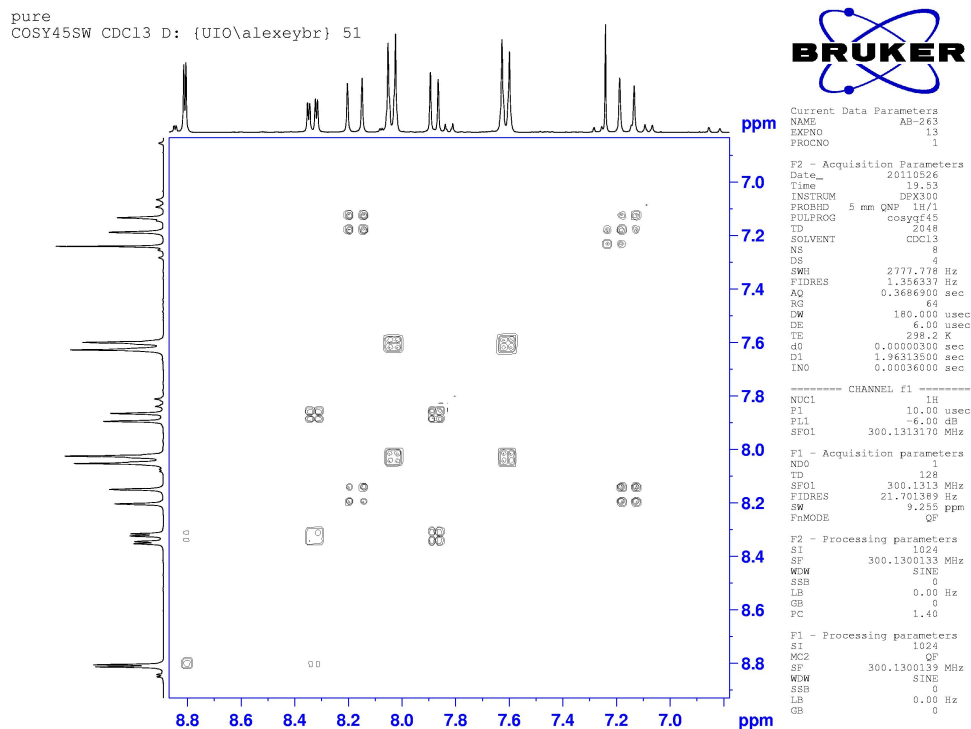
Spectrum 9, expansion of the aromatic region from spectrum 8



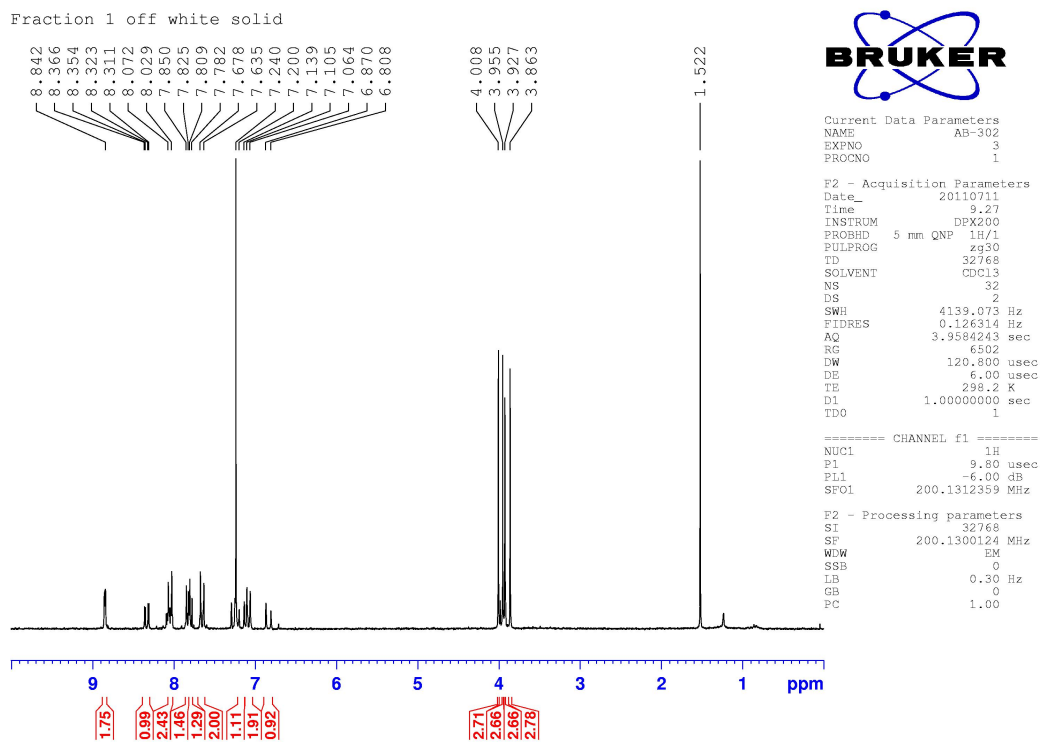
Spectrum 10



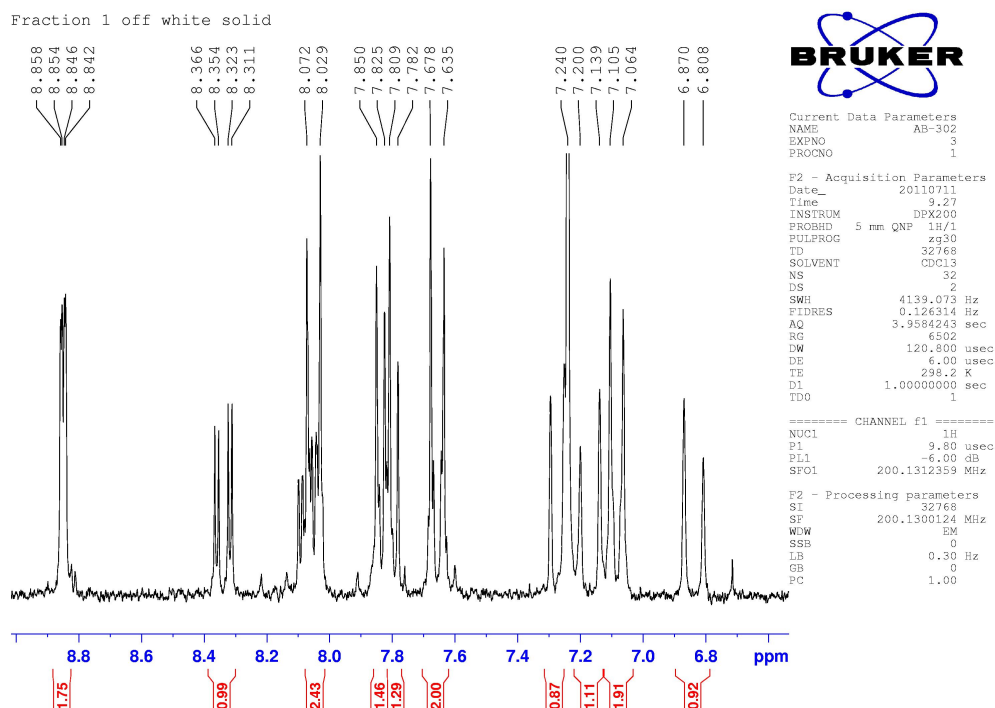
Spectrum 11



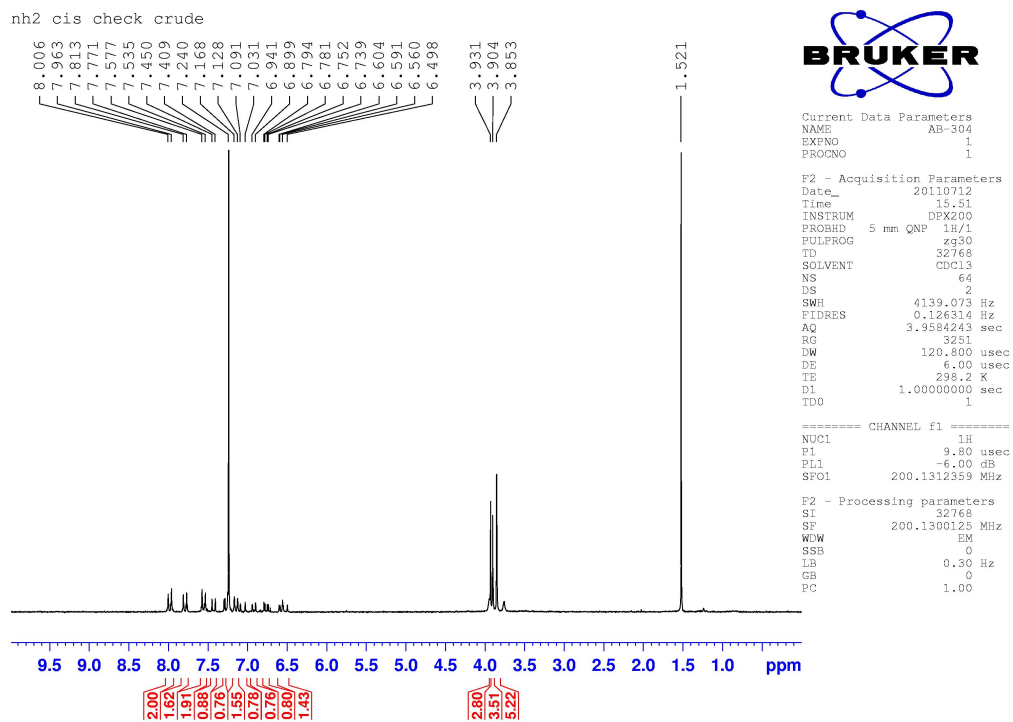
Spectrum 12



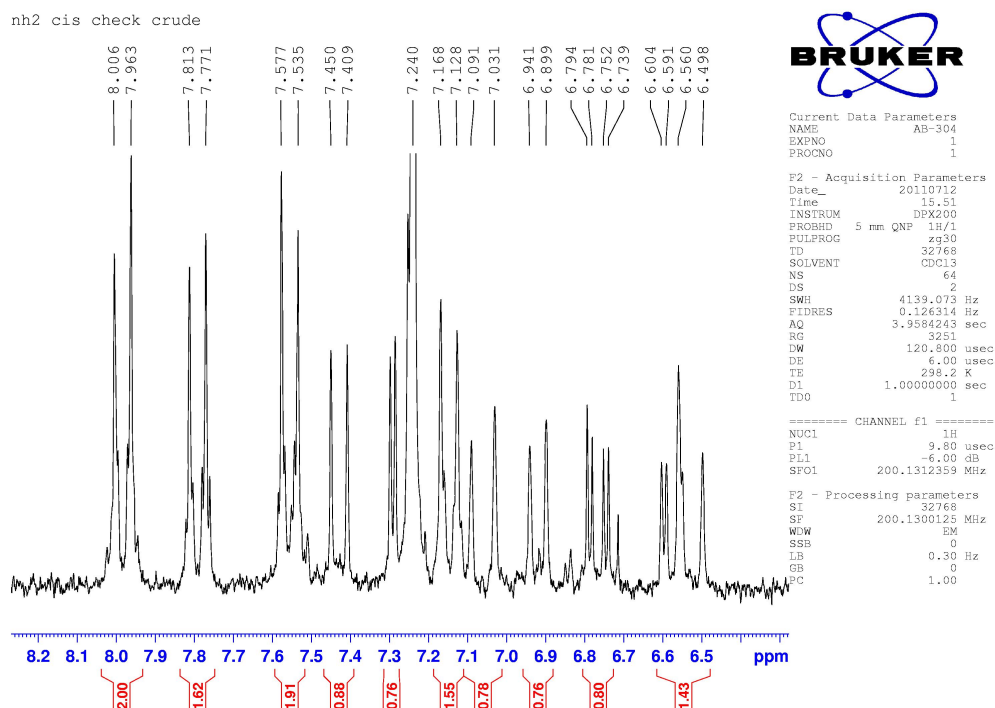
Spectrum 13, expansion of the aromatic region from spectrum 12



Spectrum 14



Spectrum 15, expansion of the aromatic region from spectrum 14



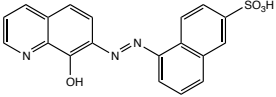
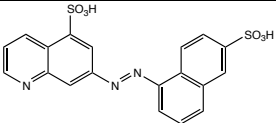
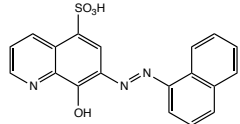
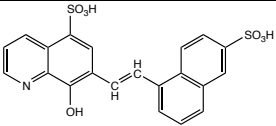
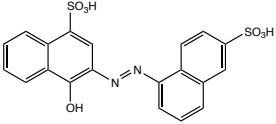
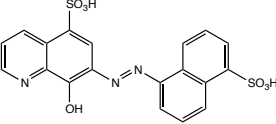
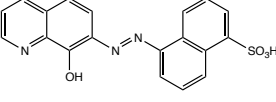
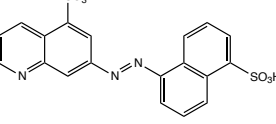
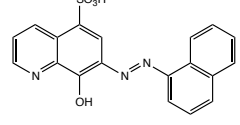
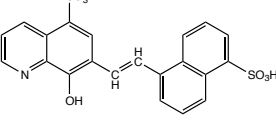
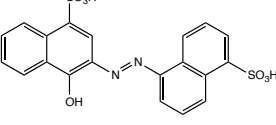
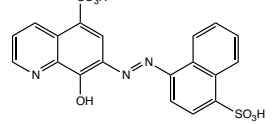
Attachment 3

Complete docking tables

Table 11. Results from the docking of compounds AB74 – AB 81-5 against the R2 region in the DBD of c-Myb

Compound	Chemical structure	logP	Residues involv. in H bonds	Fr. en. of bind. (kcal/mol)	Ki	MW	No. of H-bond accept.	No. of H-bond donors
NSC87877 / AB 81		- 0.807	Lys128 Gln129 Arg133	- 9.91	53.99 nM	459	10	3
AB 81-1		1.97	Lys128 Asn136	- 8.45	645.09 nM	379	7	2
AB 81-2		-0.516	Lys128 Arg133	- 9.25	165.66 nM	443	9	2
AB 81-3		2.225	Gly93	- 8.47	620.97 nM	379	7	2
AB 81-4		- 0.62	Lys128 Gln129 Arg133	- 9.99	47.68 nM	457	8	3
AB 81-5		0.093	Lys128 Arg133 Gln129	- 9.68	79.66 nM	458	9	3
AB 74		- 0.831	Ser116 Lys128 Arg131	- 10.94	9.5 nM	459	10	3
AB 74-1		1.946	Lys128 Arg131	- 8.91	296.19 nM	379	7	2

AB 74-2		- 0.54	Ser116 Lys128 Arg131	- 10.27	29.84 nM	443	9	2
AB 74-3		2.201	Ser116 Arg131	- 8.47	622.47 nM	379	7	2
AB 74-4		- 0.824	Ser116 Arg131	- 10.6	17 nM	457	8	3
AB 74-5		0.069	Ser116 Arg131	- 10.85	11.12 nM	458	9	3
AB 75		- 0.831	Lys128 Arg133 Asn136	- 9.08	221.08 nM	459	10	3
AB 75-1		1.946	Lys128	- 8.97	266.27 nM	379	7	2
AB 75-2		- 0.54	Lys128 Lys171	- 8.68	436.7 nM	443	9	2
AB 75-3		2.225	Lys128 Gln129	- 8.6	497.86 nM	379	7	2
AB 75-4		- 0.644	Lys128 Arg133 Asn136	- 9.31	151.1 nM	457	8	3
AB 75-5		0.069	Lys128 Arg133	- 9.18	187.14 nM	458	9	3
AB 76		- 0.831	Ser116 Lys128	- 9.32	148.38 nM	459	10	3

AB 76-1		1.946	Lys128 Arg131	- 8.38	715.6 nM	379	7	2
AB 76-2		- 0.54	Ser116 Lys128	- 9.78	68.02 nM	443	9	2
AB 76-3		2.201	Ser116 Arg131	- 8.31	805.19 nM	379	7	2
AB 76-4		- 0.824	Ser116 Lys128	- 10.03	44.45 nM	457	8	3
AB 76-5		0.069	Ser116 Lys128 Arg131	- 10.03	44.38 nM	458	9	3
AB 77		- 0.855	Glu132 Lys128 Arg133	- 9.9	55.67 nM	459	10	3
AB 77-1		1.922	Asn136 Lys128	- 7.74	2.14 μM	379	7	2
AB 77-2		- 0.564	Lys128 Arg133	- 8.68	437.63 nM	443	9	2
AB 77-3		2.201	Ser116 Lys113	- 8.12	1.11 μM	379	7	2
AB 77-4		- 0.848	Ser116 Arg131	- 8.96	268.58 nM	457	8	3
AB 77-5		0.045	Lys128 Arg133	- 9.95	50.75 nM	458	9	3
AB 78		- 0.855	Lys128 Gln129 Arg133	- 9.97	48.92 nM	459	10	3

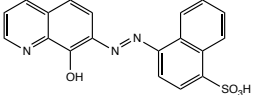
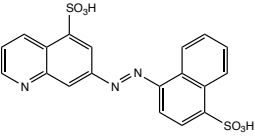
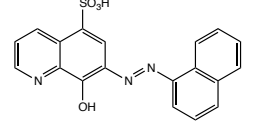
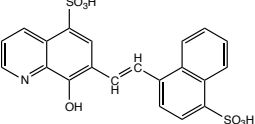
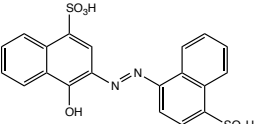
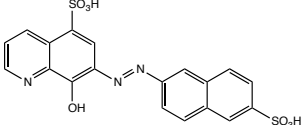
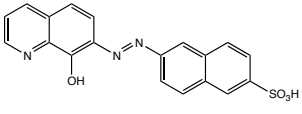
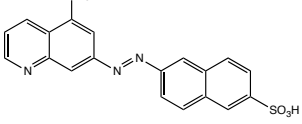
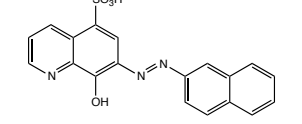
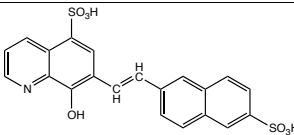
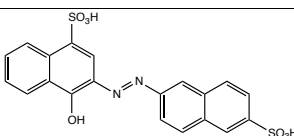
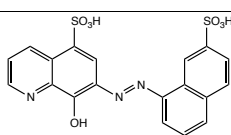
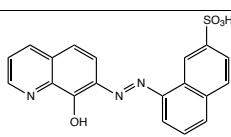
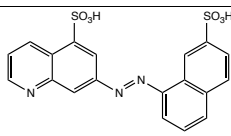
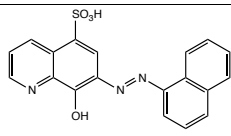
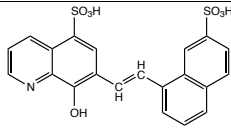
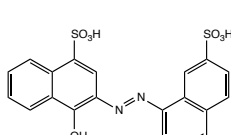
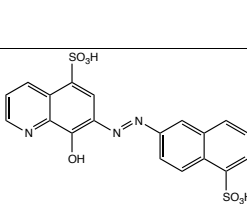
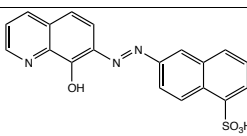
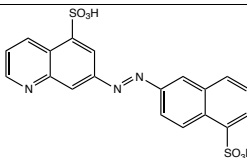
AB 78-1		1.922	Lys128	- 8.15	1.07 μ M	379	7	2
AB 78-2		- 0.564	Lys128 Arg133	- 9.12	204.84 nM	443	9	2
AB 78-3		2.201	Ser116 Arg131	- 8.29	835.69 nM	379	7	2
AB 78-4		- 0.848	Lys128 Gln129 Arg133	- 10.05	43.27 nM	457	8	3
AB 78-5		0.045	Lys128 Arg133 Gln129	-10.04	43.45 nM	458	9	3

Table 12. Results from the docking of compounds AB74 – AB81-5 against the R3 region in the DBD of c-Myb

Compound	Chemical structure	logP	Residues involv. in H bonds	Fr. en. of bind. (kcal/mol)	Ki	MW	No. of H- bond accept.	No. of H- bond donors
NSC87877 / AB 81		- 0.807	Asn179 Thr188 Asn183	- 6.63	13.73 μ M	459	10	3
AB 81-1		1.97	Asn183 His184	- 7.75	2.08 μ M	379	7	2
AB 81-2		-0.516	Thr188 Asn179 Asn183	- 6.74	11.56 μ M	443	9	2
AB 81-3		2.225	Trp147 Asn183	- 6.83	9.87 μ M	379	7	2

AB 81-4		- 0.62	Thr188	-6.9	8.69 μ M	457	8	3
AB 81-5		0.093	Asn179 Thr188 Asn183	- 6.61	14.27 μ M	458	9	3
AB 74		- 0.831	Ser187 Asn183 Ala180	-7.23	5 μ M	459	10	3
AB 74-1		1.946	Asn186 Ser187 Asn183	- 7.06	6.73 μ M	379	7	2
AB 74-2		- 0.54	His184 Asn186 Asn183	- 7.09	6.36 μ M	443	9	2
AB 74-3		2.201	NO BONDS	- 6.76	11.09 μ M	379	7	2
AB 74-4		- 0.824	Asn183 Arg190 Asn186	- 7.11	6.09 μ M	457	8	3
AB 74-5		0.069	Asn183 Ala180 Asn186 Ser187	- 7.04	6.86 μ M	458	9	3
AB 75		- 0.831	Trp147 Arg190 Ala180	- 6.49	17.35 μ M	459	10	3
AB 75-1		1.946	Thr188 His184 Asn183	- 7.9	1.61 μ M	379	7	2
AB 75-2		- 0.54	Arg190 Ser187 His184	- 6.29	24.68 μ M	443	9	2

AB 75-3		2.225	Trp147 Ser187	- 6.47	18.21 μ M	379	7	2
AB 75-4		- 0.644	Trp147 Arg190 Ala180	- 6.63	13.77 μ M	457	8	3
AB 75-5		0.069	Trp147 Arg190	- 6.79	10.47 μ M	458	9	3
AB 76		- 0.831	Arg190 His184	- 7.5	3.19 μ M	459	10	3
AB 76-1		1.946	Ser187 His184 Asn183	- 7.71	2.23 μ M	379	7	2
AB 76-2		- 0.54	Thr145 His184 Arg190 Ser187 Trp147	- 7.12	6.03 μ M	443	9	2
AB 76-3		2.201	Trp147 His184	- 7.11	6.12 μ M	379	7	2
AB 76-4		- 0.824	His184 Arg190	- 7.6	7.71 μ M	457	8	3
AB 76-5		0.069	His184 Arg190 Trp147	- 7.65	2.47 μ M	458	9	3
AB 77		- 0.855	Trp147 Arg190 His184	- 7.85	1.77 μ M	459	10	3
AB 77-1		1.922	His184 Ser187	- 7.19	5.14 μ M	379	7	2

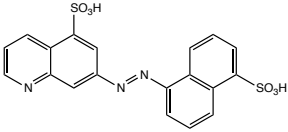
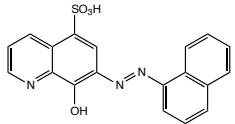
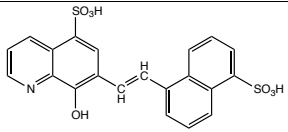
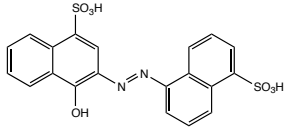
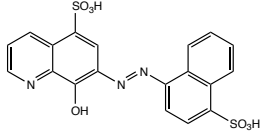
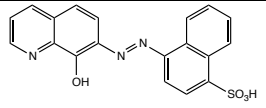
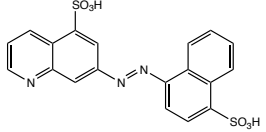
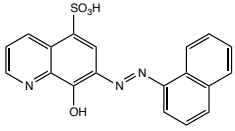
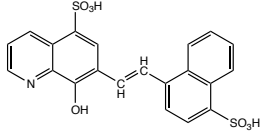
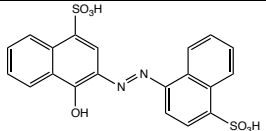
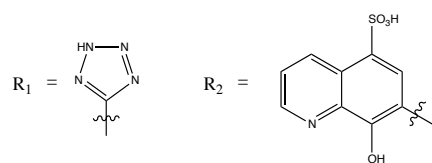
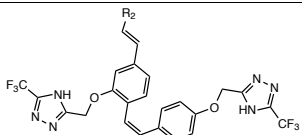
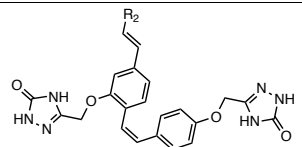
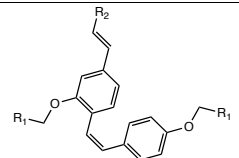
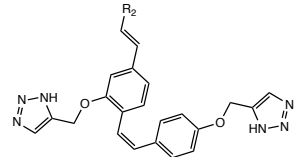
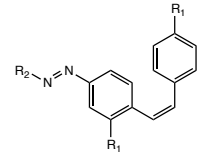
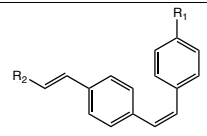
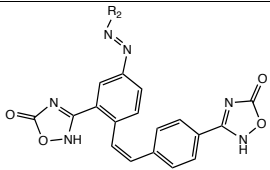
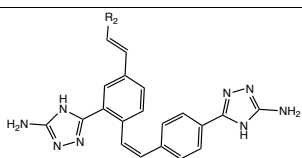
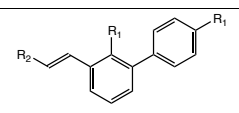
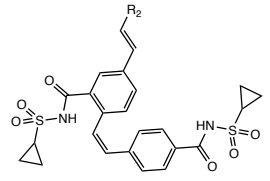
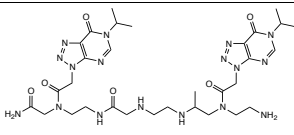
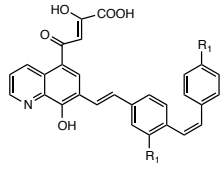
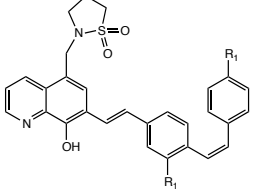
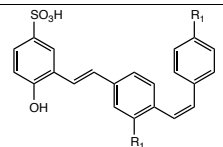
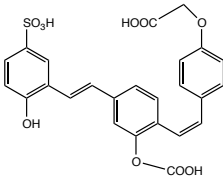
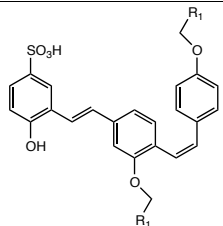
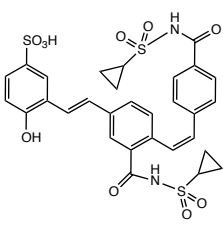
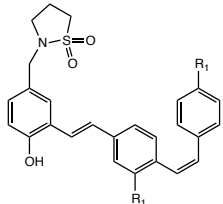
			Asn183						
			Thr177						
AB 77-2		- 0.564	Ser187 Arg190 His184 Trp147	- 6.88	9 μM	443	9	2	
AB 77-3		2.201	Asn183	- 6.87	9.2 μM	379	7	2	
AB 77-4		- 0.848	Trp147 Asn183	- 6.87	9.17 μM	457	8	3	
AB 77-5		0.045	Arg190 Trp147 His184	- 7.84	1.8 μM	458	9	3	
AB 78		- 0.855	Arg190 Trp147	- 6.81	10.24 μM	459	10	3	
AB 78-1		1.922	Asn183 Ser187	- 7.64	2.52 μM	379	7	2	
AB 78-2		- 0.564	Thr188 Asn179 His184	- 7.71	2.21 μM	443	9	2	
AB 78-3		2.201	Asn183 His184 Trp147	- 7.21	7.35 μM	379	7	2	
AB 78-4		- 0.848	Arg190 Ala180 Trp147	- 7.03	6.99 μM	457	8	3	
AB 78-5		0.045	Arg190 Trp147 Ala180	- 6.93	8.36 μM	458	9	3	

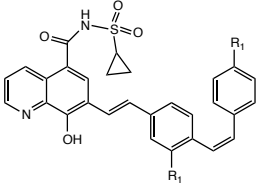
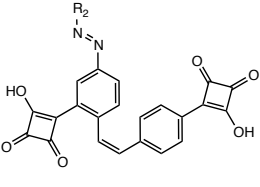
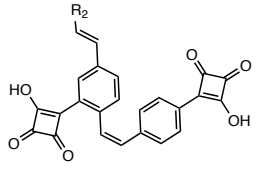
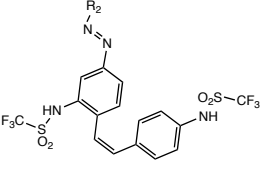
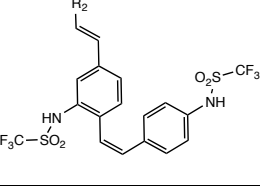
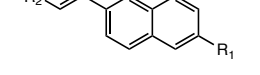
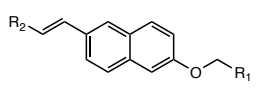
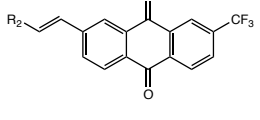
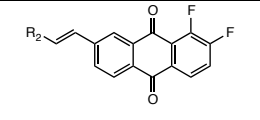
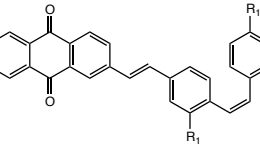
Table 13. Results from the docking of compounds ABAD1 – ABAD68 against the R2 region in the DBD of c-Myb

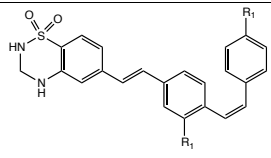
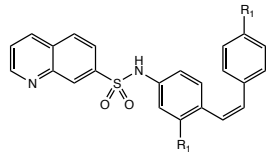
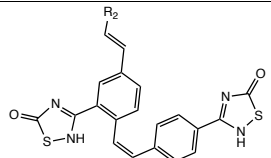
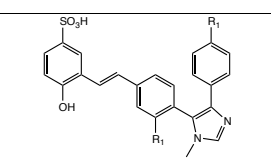
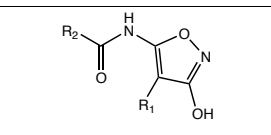
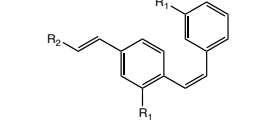
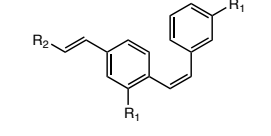
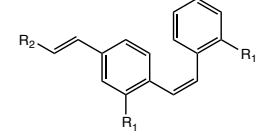
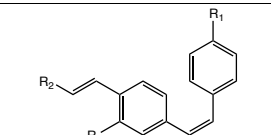
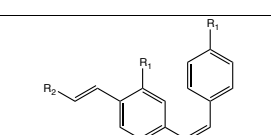
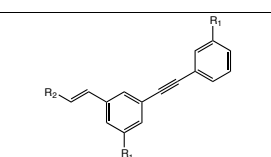


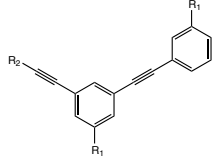
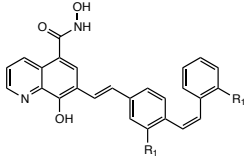
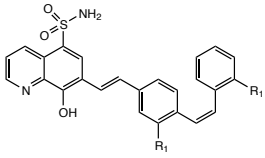
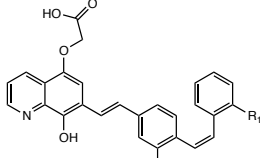
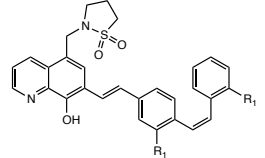
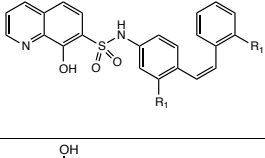
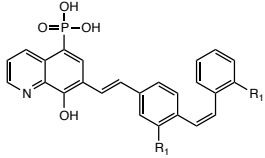
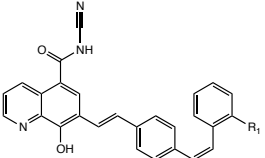
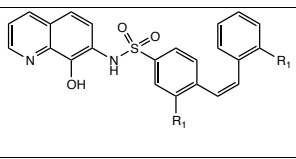
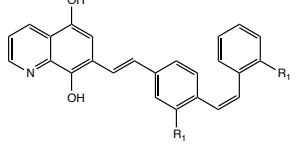
Compound	Chemical structure	logP	Residues involv. in H bonds	Fr. en. of bind. (kcal/mol)	Ki	MW	No. of H- bond accept.	No. of H- bond donors
ABAD 1		3.292	Trp115	- 10.83	11.43 nM	564	12	4
ABAD2		2.393	No bonds	- 11.79	2.26 nM	565	13	4
ABAD3		2.165	Lys128 Arg131	-8.17	1.03 μM	659	11	4
ABAD4		2.14	Ser116 Lys128 His135	- 10.22	32.19 nM	577	11	4
ABAD5		1.93	Gln129	- 7.98	1.41 μM	687	13	6
ABAD6		1.957	Glu132	- 8.07	1.22 μM	623	13	4
ABAD7		1.564	Glu132	- 5.58	81.62 μM	653	15	8

ABAD8		3.958	Arg125 Lys128	- 8.1	1.15 μM	759	13	4
ABAD9		1.245	Lys128 Gln129 Lys171	- 8.01	1.35 μM	655	15	6
ABAD10		1.675	Glu132 Ser116	- 11.2	6.17 nM	625	15	4
ABAD11		3.012	Trp95	- 9.07	226.45 nM	623	13	4
ABAD12		1.488	No hydr bonds	- 10.78	12.52 nM	627	17	4
ABAD13		3.216	Lys113 Glu132	- 7.41	3.67 μM	497	9	3
ABAD14		1.983	Ser116 Arg131	- 8.67	438.4 nM	599	15	4
ABAD15		2.282	Glu132 Lys128	- 7.02	7.16 μM	593	13	8
ABAD16		1.606	Lys128 Arg131	- 9.17	190.88 nM	539	13	4
ABAD18		1.403	Lys128	- 10.35	26.06 nM	723	12	4

ABAD19		-	Glu132 Pro174 Trp115	- 3.8	1.63 mM	-	--	-
ABAD20		4.355	Glu132 Arg131 Ser116	- 8.25	903.44 nM	599	14	5
ABAD21		4.712	Glu132 Ser116	- 10.84	11.32 nM	618	13	3
ABAD23		1.685	No bonds	- 9.61	89.95 nM	516	12	4
ABAD24		1.433	Lys128 His135 Ser116	- 9.55	100.22 nM	528	10	4
ABAD25		0.967	Ser116 Lys128 His135 Trp115	- 8.83	335 nM	576	14	4
ABAD26		0.787	Gly127 Lys128 Glu132 His135 Trp115	- 7.17	5.58 μM	688	12	4
ABAD27		4.518	Gly127 Lys128	- 9.51	107 nM	555	12	3

ABAD28		5.234	Arg131	- 10.24	31.34 nM	631	13	4
ABAD29		2.178	Ser116 His135 Lys128 Arg131	- 8.92	290 nM	623	13	4
ABAD30		2.365	Glu132 His135	- 8.61	491 nM	621	11	4
ABAD31		8.647	Lys113 Asn136	- 8.7	420 nM	725	13	4
ABAD32		8.741	Arg 131	- 9.87	58.17 nM	723	11	4
ABAD33		1.817	Lys128 Arg133	-8.89	306 nM	445	9	3
ABAD34		1.458	Asn136 Arg133 Ala167 Lys128	- 9.3	151.4 nM	475	10	3
ABAD35		3.807	Lys128 Asn136	- 7.66	2.43 μM	525	7	2
ABAD36		3.167	Arg133 Asn136	- 7.91	1.6 μM	493	7	2
ABAD37		7.367	Asn136 Arg131	- 8.76	379.6 nM	548	10	2

ABAD38		4.363	No bonds	- 8.49	602.21 nM	524.57	12	4
ABAD39		3.848	Lys128 Arg131 Trp115	- 10.38	24.57 nM	522	12	3
ABAD40		3.453	No bonds	- 6.57	15.25 μM	629	11	4
ABAD41		1.412	No bonds	- 6.29	24.2 μM	568	14	4
ABAD42		- 1.842	Arg 133 Asn 136	- 8.83	338.03 nM	419	14	9
ABAD43		2.369	Trp115	- 9.13	202.59 nM	565	13	4
ABAD44		2.369	Gln129 Lys128	- 10.87	10.82 nM	565	13	4
ABAD45		2.369	Gln129 Arg133	- 9.84	61.63 nM	565	13	4
ABAD46		2.393	Ser116	- 10.35	25.7 nM	565	13	4
ABAD47		2.393	Ser116 Gly127	- 10.09	40.51 nM	565	13	4
ABAD48		1.419	Ser116 Arg131 Lys128	-10.01	45.79 nM	563	13	4

ABAD49		0.468	Lys128 Arg131 Ser116	- 6.68	12.59 μ M	561	13	4
ABAD50		4.05	Ser116	- 8.5	301.42 nM	544	13	5
ABAD51		3.867	Ser116	-10.41	23.31 nM	564	13	5
ABAD52		4.476	Trp115 Ser116	- 8.52	573.08 nM	559	13	4
ABAD53		4.484	Ser116	- 10.47	21.29 nM	618	13	3
ABAD54		3.33	Arg133	- 12.87	367.91 pM	538	13	4
ABAD55		3.988	Lys128	- 10.04	43.46 nM	565	13	5
ABAD56		4.139	Trp115	- 8.58	511.22 nM	553	13	4
ABAD57		3.33	No bonds	- 9.7	77.94 nM	538	13	4
ABAD58		4.95	Ser 116 Lys 113 Glu 132	- 9.25	167.2 nM	501	11	4

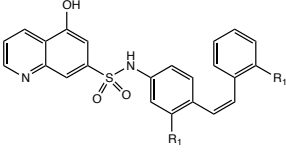
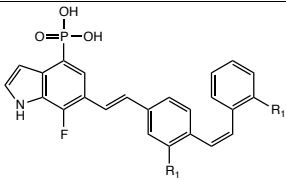
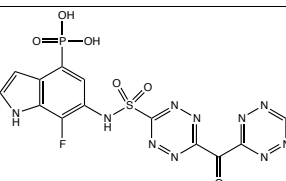
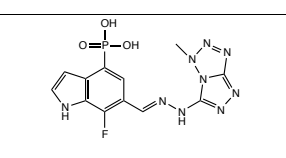
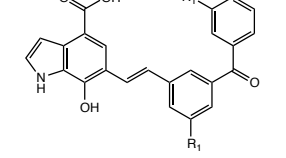
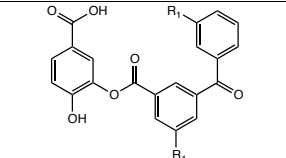
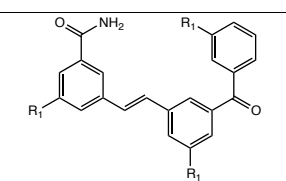
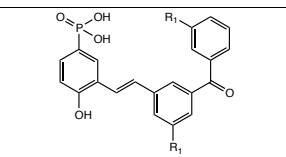
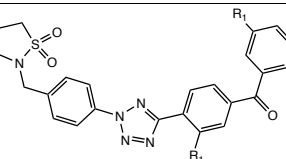
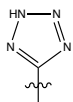
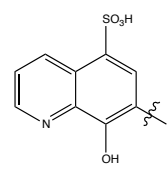
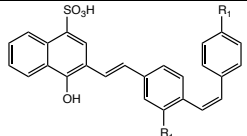
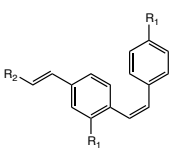
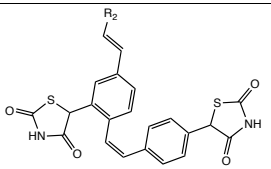
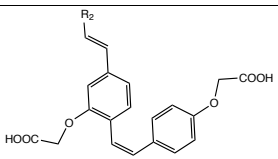
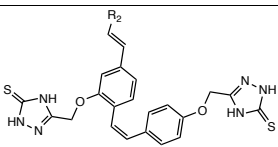
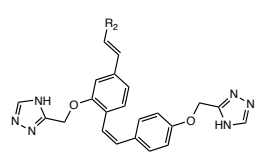
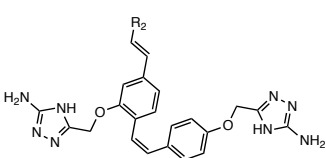
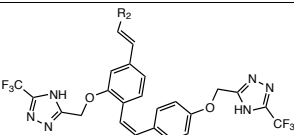
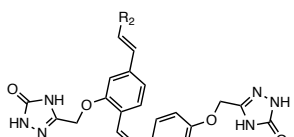
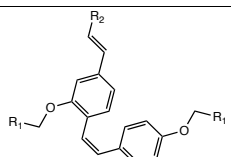
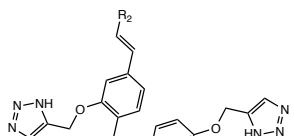
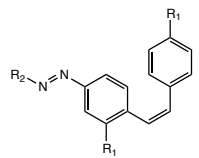
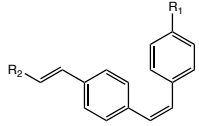
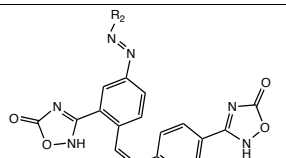
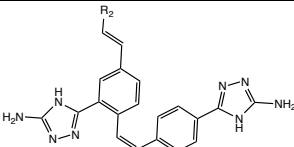
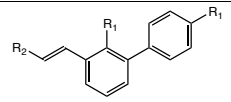
ABAD59		3.63	Gln 129 Glu132	- 10.94	9.6 nM	538	13	4
ABAD60		4.29	Ser 116	- 7.79	1.95 μM	555	12	5
ABAD61		-1.67	Trp 95 Arg125 Arg133	- 7.4	3.77 μM	482	16	4
ABAD62		-0.21	Trp115 His135	-4.84	282.95 μM	379	12	4
ABAD63		3.94	Glu132 Ser116 Arg131	-10.27	29.69 nM	519	13	5
ABAD64		3.21	Gly93 Trp95 Lys171	-9.57	96.35 nM	498	14	4
ABAD65		2.88	Ser116 Gly127	-10.31	27.87 nM	531	15	5
ABAD67		3.27	Ser116	-8.12	1.12 μM	516	13	5
ABAD68		2.24	Ser116 Arg131	-8.16	1.05 μM	595	16	2

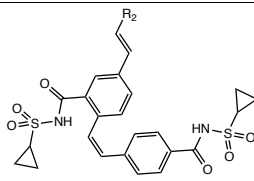
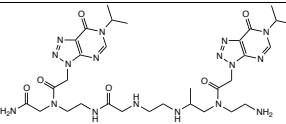
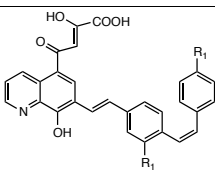
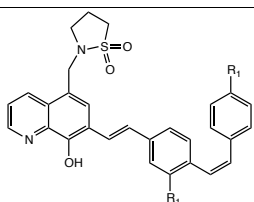
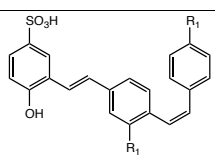
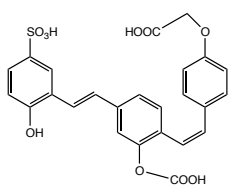
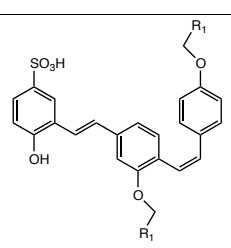
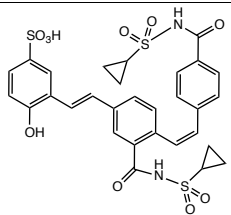
Table 14. Results from the docking of compounds ABAD1 – ABAD 68 against the R3 region in the DBD of c-Myb

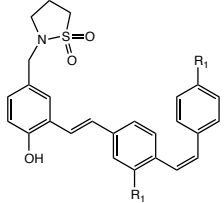
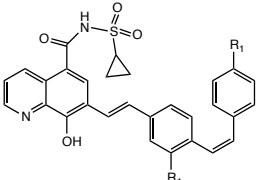
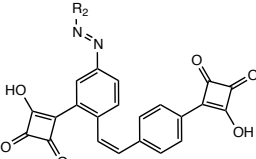
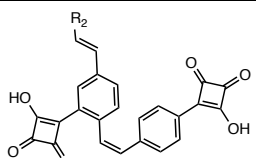
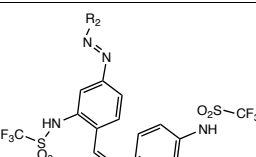
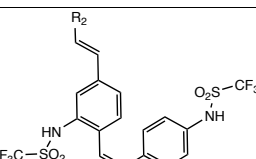
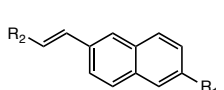
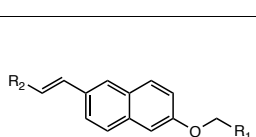
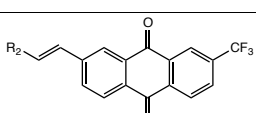
$R_1 =$


$R_2 =$


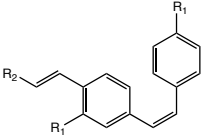
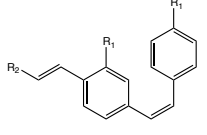
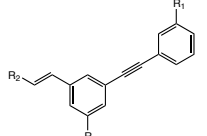
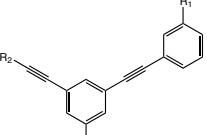
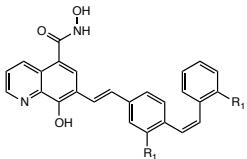
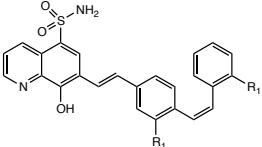
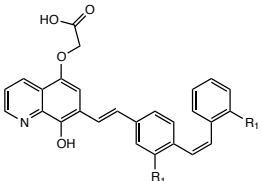
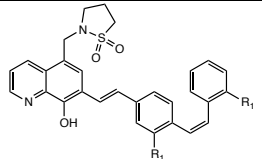
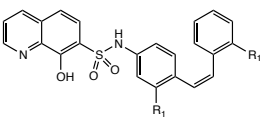
Compound	Chemical structure	logP	Residues involv. in H bonds	Fr. en. of bind. (kcal/mol)	Ki	MW	No. of H- bond accept.	No. of H- bond donors
ABAD 1		3.292	Arg190 Asn183 Asn186	- 10.16	35.98 nM	564	12	4
ABAD2		2.393	Asn186 Arg190 Thr177 Ala180	- 10.54	18.64 nM	565	13	4
ABAD3		2.165	Ala180	- 7.76	2.05 μM	659	11	4
ABAD4		2.14	Arg190 Arg191 Ser187	- 7.53	3 μM	577	11	4
ABAD5		1.93	Asn183 Trp147 Arg190	- 8.21	952.32 nM	687	13	6
ABAD6		1.957	Asn186 Arg190 Thr177 Thr188	- 6.14	31.65 μM	623	13	4
ABAD7		1.564	Asn139 Pro140 Asn183 Ser187	- 5.21	150.7 μM	653	15	8

ABAD8		3.958	Asn186	- 6.54	15.99 μ M	759	13	4
ABAD9		1.245	Ser187 Arg191 Asn183 Trp147	- 6.79	10.46 μ M	655	15	6
ABAD10		1.675	Arg190 Trp147	- 10.26	30.21 nM	625	15	4
ABAD11		3.012	Asn179 Asn183 Lys182 Trp147 Asn183	- 5.36	118 μ M	623	13	4
ABAD12		1.488	Ala180 Asn186 Asn183 Arg190	-10.56	18.06 nM	627	17	4
ABAD13		3.216	His184 Thr188 Arg190	- 6.38	21.21 μ M	497	9	3
ABAD14		1.983	Asn183 His184	- 8.33	786.81 nM	599	15	4
ABAD15		2.282	Ala180 Arg190	- 6.35	22.34 μ M	593	13	8
ABAD16		1.606	His137 Asn183	- 6.97	7.76 μ M	539	13	4

ABAD18		1.403	His184 Arg191 Trp147	- 7.41	3.68 μ M	723	12	4
ABAD19		-	Ser187 Asn186	- 4.1	992.9 μ M	-	--	-
ABAD20		4.355	No bonds	- 7.83	1.82 μ M	599	14	5
ABAD21		4.712	His184	- 9.15	197.17 nM	618	13	3
ABAD23		1.685	Thr188 Asn186 Arg190	- 10.03	44.59 nM	516	12	4
ABAD24		1.433	Ala180 Trp147 Lys182 His184 Arg190	- 7.82	1.86 μ M	528	10	4
ABAD25		0.967	Arg190 Asn183 Ala180 Ser187	- 6.61	14.21 μ M	576	14	4
ABAD26		0.787	Arg190 Asn183 His184 Ser187	- 7.07	6.55 μ M	688	12	4

ABAD27		4.518	- 8.93	282.6 nM	555	12	3	Asn186
								Thr188
								Arg190
								Asn183
ABAD28		5.234	- 10.52	19.42 nM	631	13	4	Thr177
								Arg190
ABAD29		2.178	- 7.49	3.21 μM	623	13	4	Ser187
								Asn179
								Asn183
								His184
ABAD30		2.365	- 9.18	186.68 nM	621	11	4	Asn183
								Thr177
								Thr188
								His184
ABAD31		8.647	- 7.53	3.02 μM	725	13	4	Trp147
								Ala180
								Ser187
								Arg190
ABAD32		8.741	- 6.59	14.77 μM	723	11	4	Trp147
								His184
ABAD33		1.817	- 6.42	19.55 μM	445	9	3	Asn179
								Thr188
								Asn183
ABAD34		1.458	- 6.08	35.18 μM	475	10	3	Trp147
								Arg190
								Ala180
ABAD35		3.807	- 6.36	21.87 μM	525	7	2	Asn179
								Asn183

ABAD36		3.167	Asn183 His184 Asn179	- 6.01	39.27 μ M	493	7	2
ABAD37		7.367	Trp147 Asn183	- 8.54	552.15 nM	548	10	2
ABAD38		4.363	Ser187 Arg190	- 7.13	5.93 μ M	524.57	12	4
ABAD39		3.848	Trp 147 Asn186	- 8.69	426.76 nM	522	12	3
ABAD40		3.453	Asn 186	- 7.13	5.97 μ M	629	11	4
ABAD41		1.412	Arg190 Asn186 Asn183	- 5.78	57.95 μ M	568	14	4
ABAD42		- 1.842	Arg190 Asn186 Asn183	- 7.43	3.57 μ M	419	14	9
ABAD43		2.369	Asn183	- 9.89	56.41 nM	565	13	4
ABAD44		2.369	Asn183 Arg190	- 10.67	14.98 nM	565	13	4
ABAD45		2.369	Asn186 Arg190 Ser187 His184	- 11.47	3.9 nM	565	13	4

ABAD46		2.393	His184	- 6.43	19.41 μM	565	13	4
ABAD47		2.393	Ala180 Ser187 Trp147 Asn183	- 7.49	3.26 μM	565	13	4
ABAD48		1.419	Asn139 Asn186 Trp147	- 6.83	9.8 μM	563	13	4
ABAD49		0.468	Asn186 His184 Asn139	- 6.9	8.8 μM	561	13	4
ABAD50		4.05	Ala 180 Asn 186 Trp 147 Asn 183	- 9.12	207.19 nM	544	13	5
ABAD51		3.867	Arg 190	- 10.41	23.22 nM	564	13	5
ABAD52		4.476	Arg 190	- 9.97	48.82 nM	559	13	4
ABAD53		4.484	Asn 186	- 10.44	22.16 nM	618	13	3
ABAD54		3.33	Asn 183 Asn 186 Ser 187	- 9.56	98.61 nM	538	13	4

ABAD55		3.988	Asn 183 Asn 186 Ser 187	- 9.89	55.88 nM	565	13	5
ABAD56		4.139	Asn 183 Ser 187	- 8.56	532.06 nM	553	13	4
ABAD57		3.33	No bonds	- 8.65	457.09 nM	538	13	4
ABAD58		4.95	Asn 186 Thr 188 His 184	- 8.2	972.63 nM	501	11	4
ABAD59		3.63	His 184 Thr 188	- 8.43	658.44 nM	538	13	4
ABAD60		4.29	Trp 147 Asn 186 Arg 190 Ala 180	- 8.34	772.42 nM	555	12	5
ABAD61		-1.67	Trp 147 Asn 183 His 184	- 5.69	67.23 μM	482	16	4
ABAD62		-0.21	Asn183 Asn186	-4.64	398.44 μM	379	12	4
ABAD63		3.94	Asn183 Asn186	-7.14	5.85 μM	519	13	5

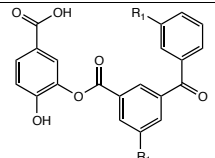
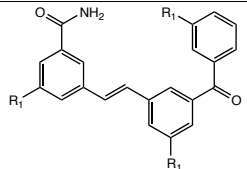
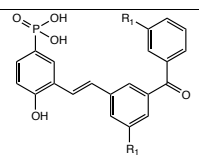
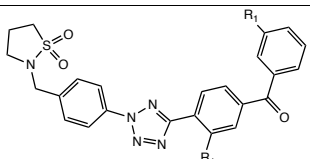
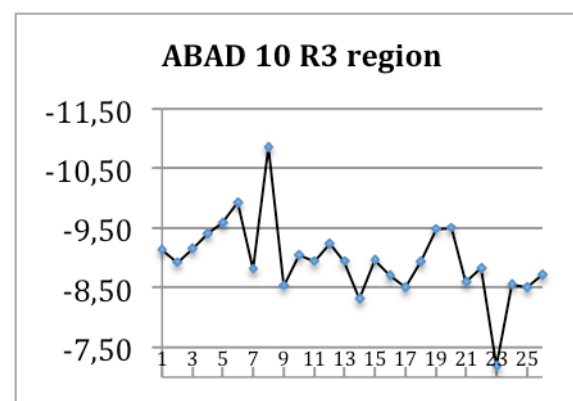
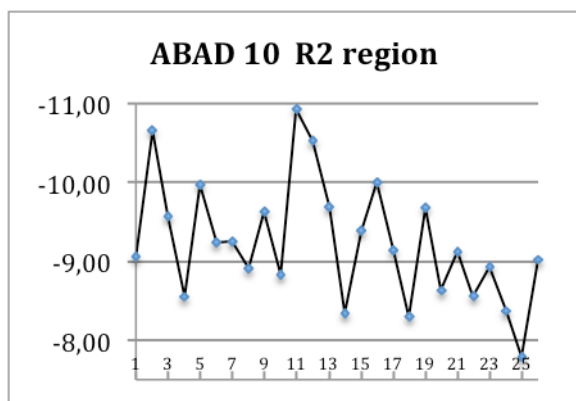
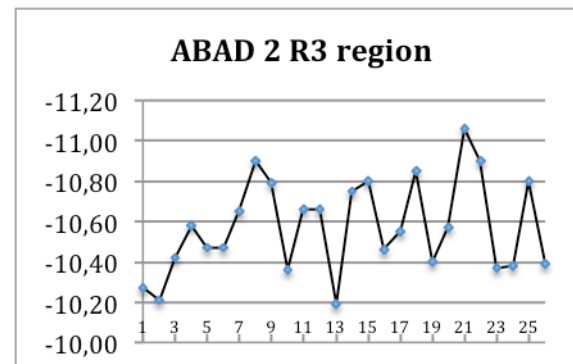
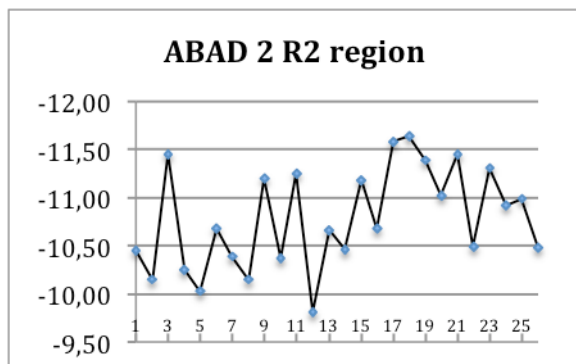
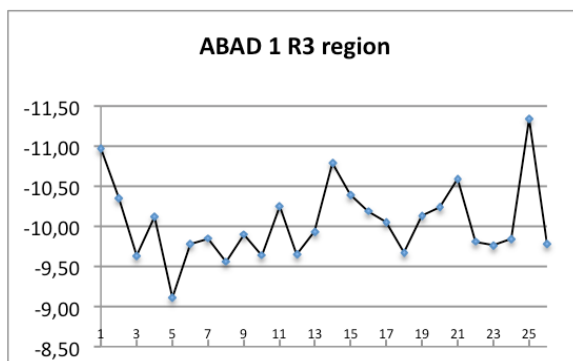
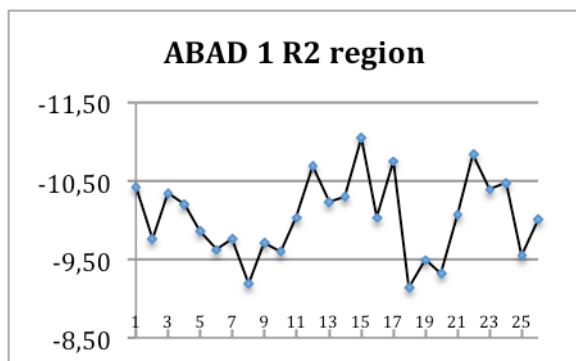
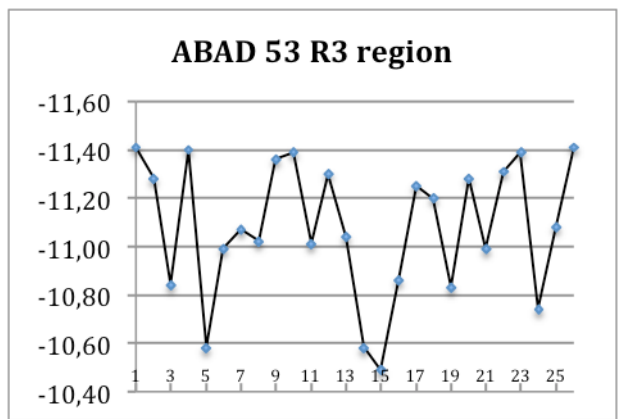
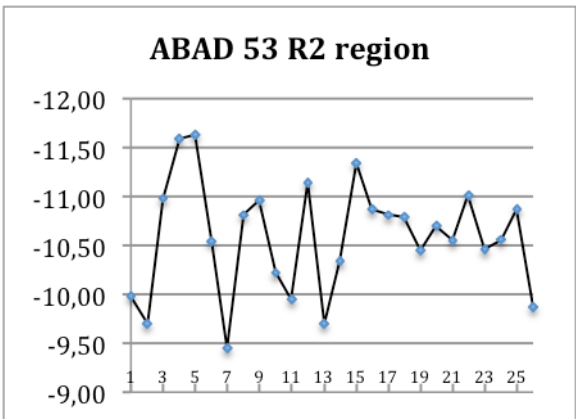
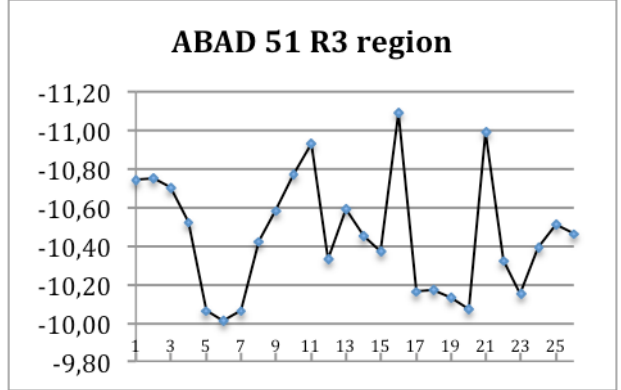
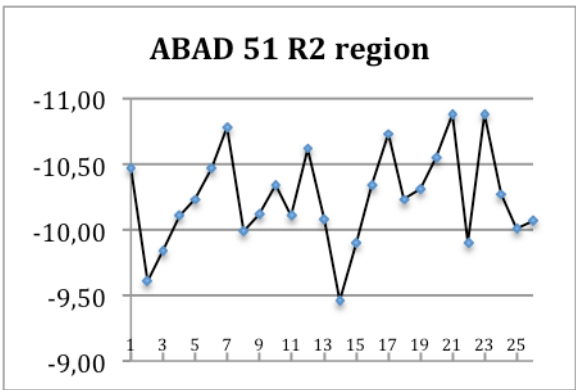
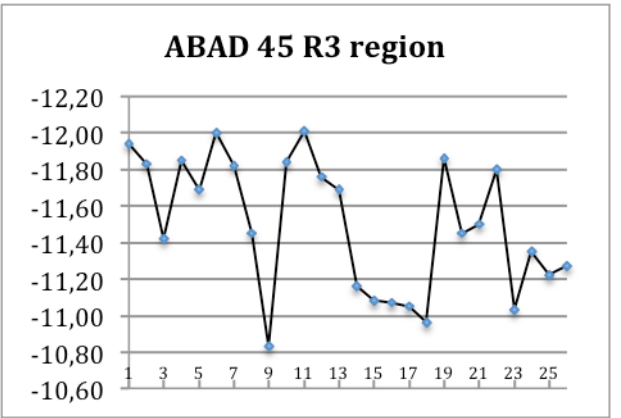
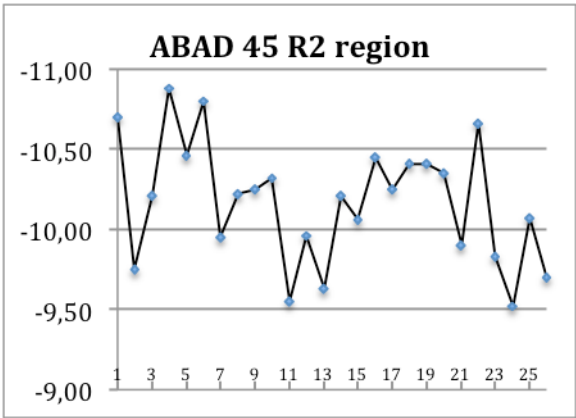
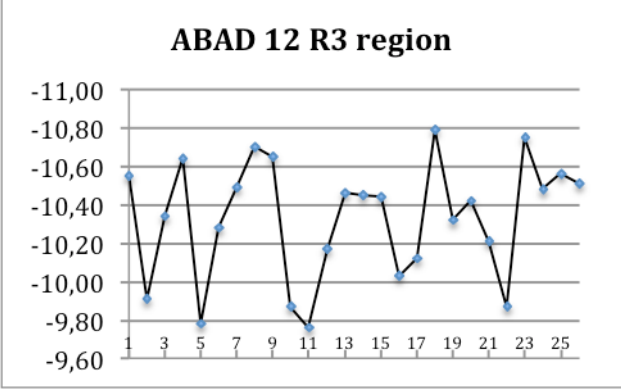
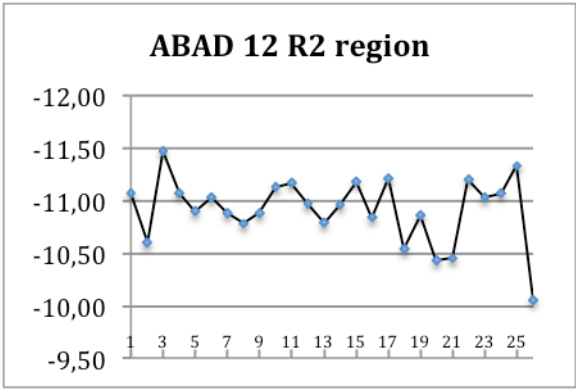
ABAD64		3.21	Asn183	-7.01	7.16 μ M	498	14	4
ABAD65		2.88	Asn183 Arg190	-7.78	1.97 μ M	531	15	5
ABAD67		3.27	Thr177 Asn186	-6.65	13.46 μ M	516	13	5
ABAD68		2.24	Trp147 Arg190	-7.82	1.87 μ M	595	16	2

Table 15. Re-docking of ABAD1, 2, 10, 12, 45, 51, and 53 against the R2 and R3 regions in the DBD of c-Myb

X-axis: Number of the run. Y-axis: The free energy of binding in kcal/mol





Attachment 4

AutoDock visualizations

Docking of AB81, 81-1, 81-2, 81-3, 81-4 and 81-5 against the R2 region of the DBD in c-Myb

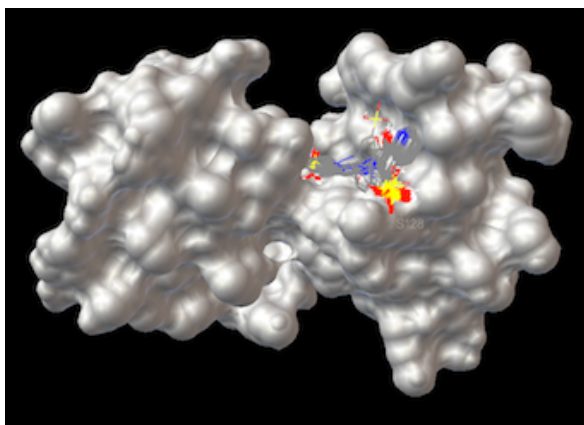


Fig. 68. Top 5 conformations of every compound. One can see that they bind in the same pocket/site.

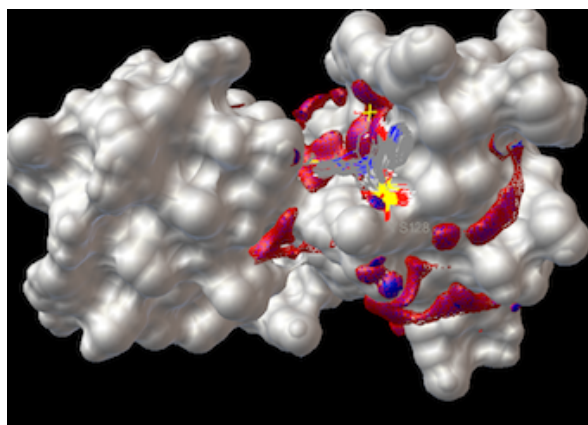


Fig. 69. Red and blue areas are sites that have affinity for nitrogen (blue) and oxygen (red). The same areas have affinity for both nitrogen and oxygen.

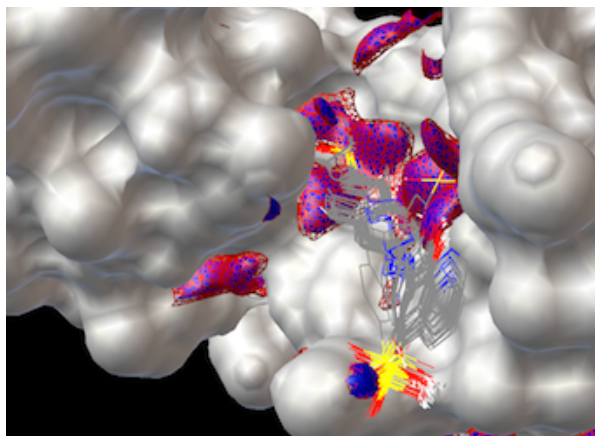


Fig. 70. Same as fig. 69, just closer look. The sulfonic acid groups in all compounds are buried or lay near the areas with affinity for oxygen.

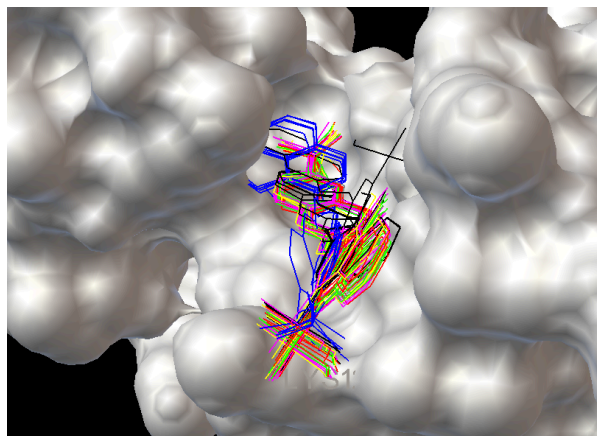


Fig. 71. AB81 – green, AB81-1 – blue, AB81-2 – red, AB81-3 – black, AB81-4 – purple, AB81-5 – yellow.

Docking of AB81, 81-1, 81-2, 81-3, 81-4 and 81-5 against the R3 region of the DBD in c-Myb

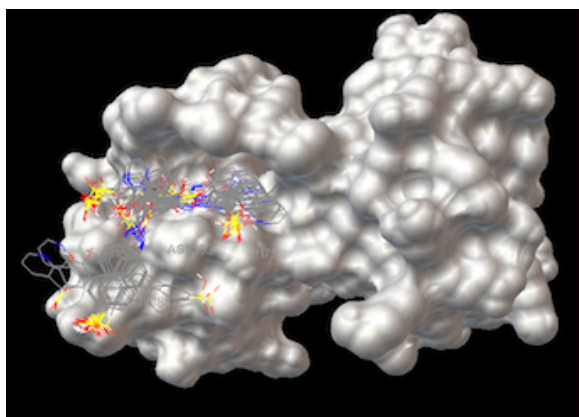


Fig. 72. Top 5 conformations of every compound. One can see that they bind in the same pocket/site.

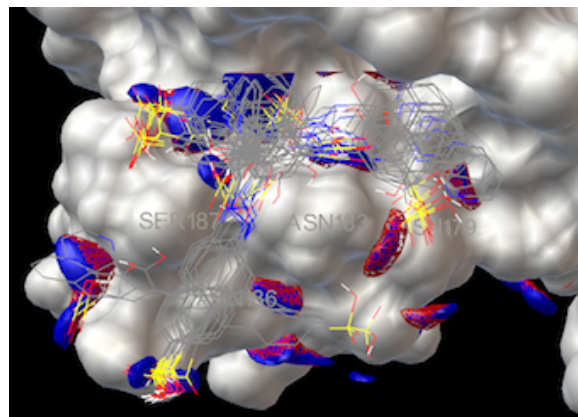


Fig. 73. Red and blue areas are sites with affinity for nitrogen and oxygen. The same areas have affinity for both nitrogen and oxygen

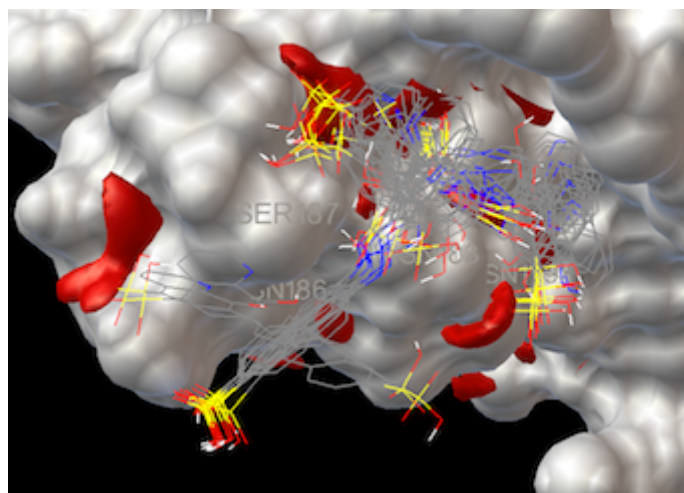


Fig. 74. Oxygen affin. map. The sulfonic acid groups are buried inside the areas with affinity for the oxygen.

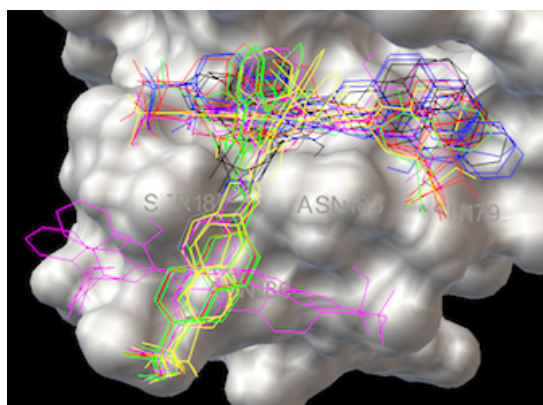


Fig. 75. AB81 – green, AB81-1 – blue, AB81-2 – red, AB81-3 – black, AB81-4 – purple, AB81-5 – yellow.

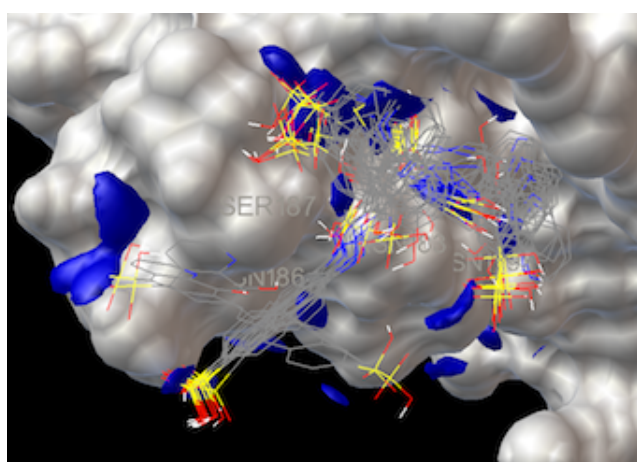


Fig. 76. Nitrogen affinity map.

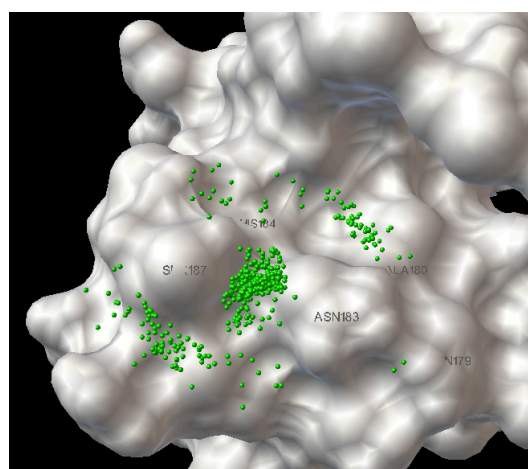


Fig. 77. All 600 conformations represented with green spheres.

Docking of AB74, 74-1 – 74-5 against the R2 region of the DBD in c-Myb

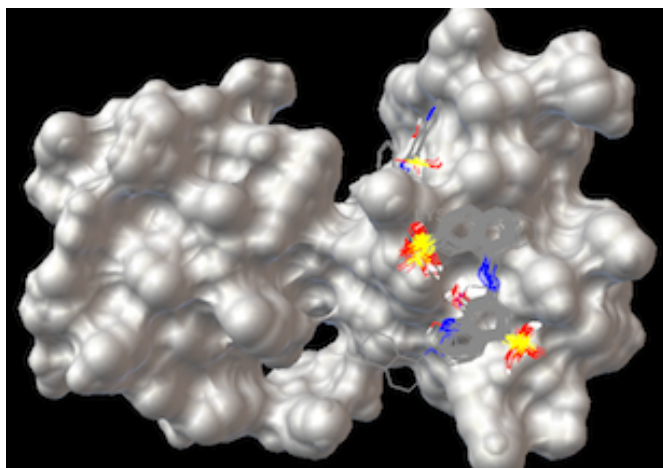


Fig. 78. Top 5 conformations of every compound. One can see that they bind in the same pocket/site.

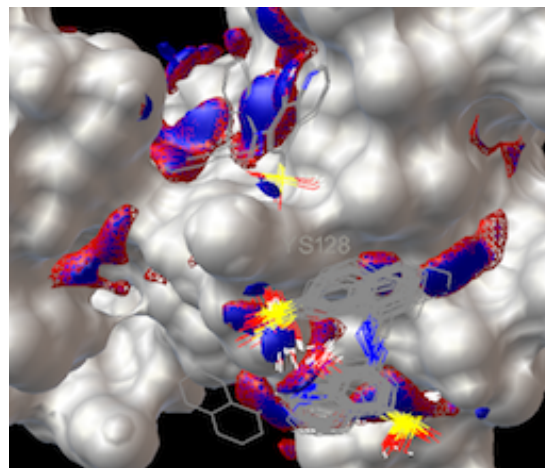


Fig. 79. Red and blue areas are sites with affinity for nitrogen and oxygen.

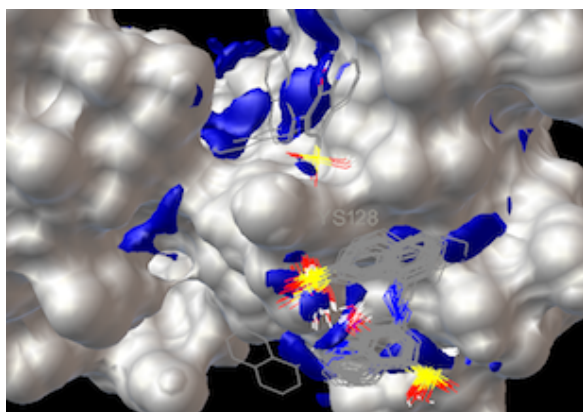


Fig. 80. Nitrogen affinity map.

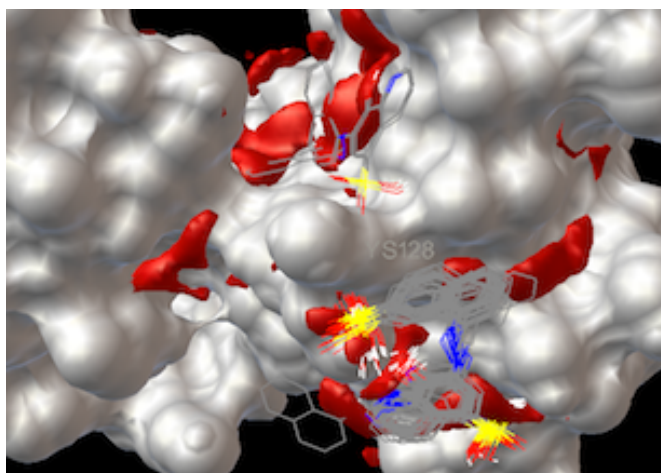


Fig. 81. Oxygen affinity map. The sulfonic acid groups are buried inside the areas with affinity for oxygen.

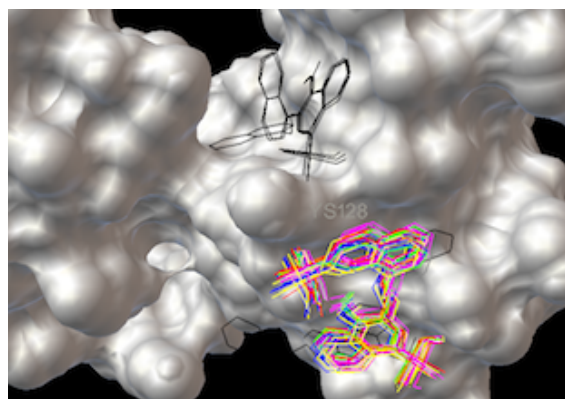


Fig. 82. AB74 – green, AB74-1 – blue, AB74-2 – red, AB74-3 – black, AB74-4 – purple, AB74-5 – yellow.

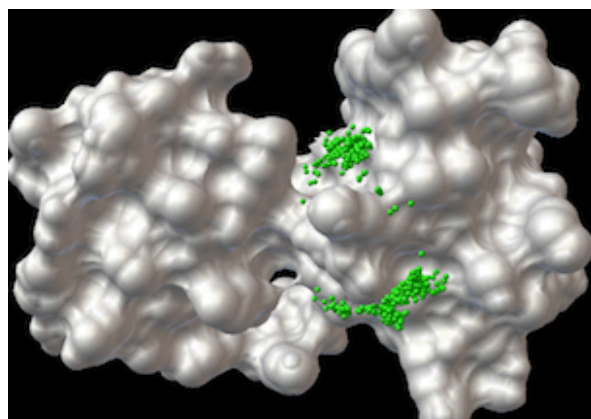


Fig. 83. All 600 conformations represented with green spheres.

Docking of AB74, 74-1 – 74-5 against the R3 region of the DBD in c-Myb

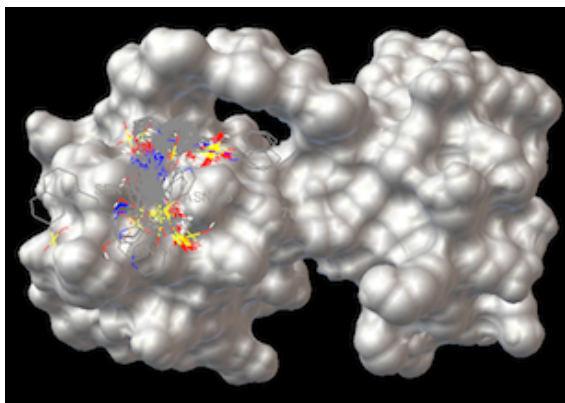


Fig. 84. Top 5 conformations of every compound. One can see that they bind in the same pocket/site.

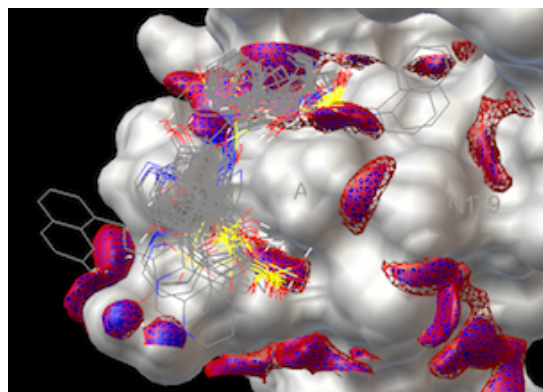


Fig. 85. Red and blue areas are sites with affinity for nitrogen and oxygen.

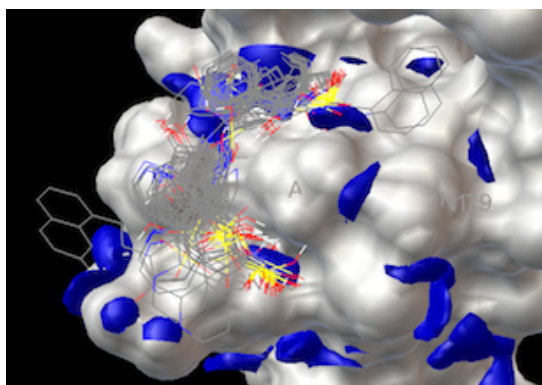


Fig. 86. Nitrogen affinity map.

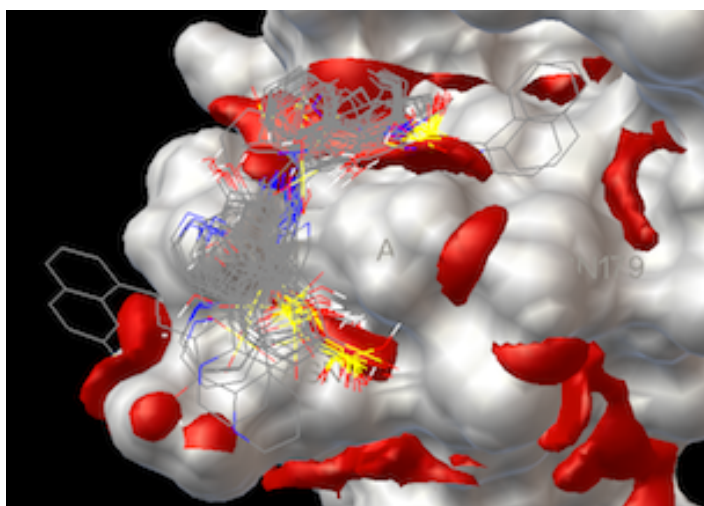


Fig. 87. Oxygen affinity map. The sulfonic acid groups are buried inside the areas with affinity for oxygen.

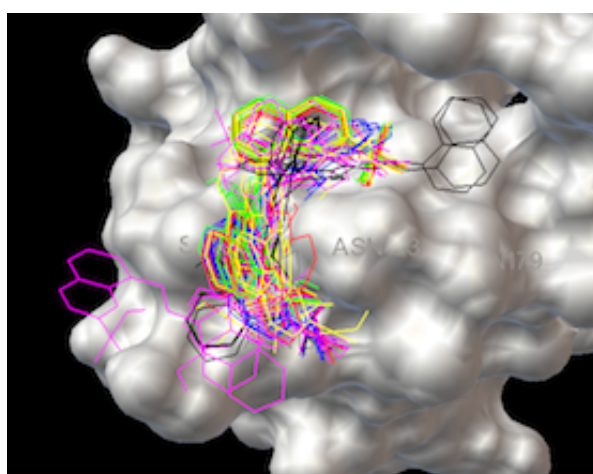


Fig. 88. AB74 – green, AB74-1 – blue, AB74-2 – red, AB74-3 – black, AB74-4 – purple, AB74-5 – yellow.

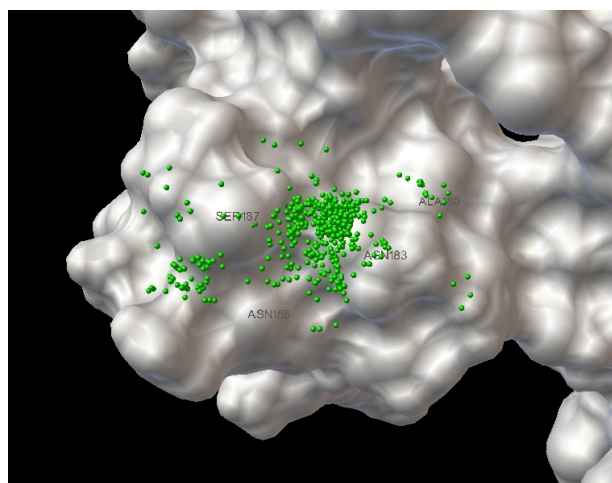


Fig. 89. All 600 conformations represented with green spheres.

Docking of AB75, 75-1 – 75-5 against the R2 region of the DBD in c-Myb

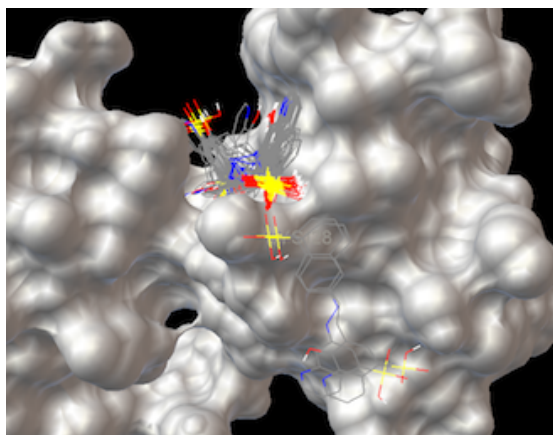


Fig. 90. Top 5 conformations of every compound. One can see that they bind in the same pocket/site.

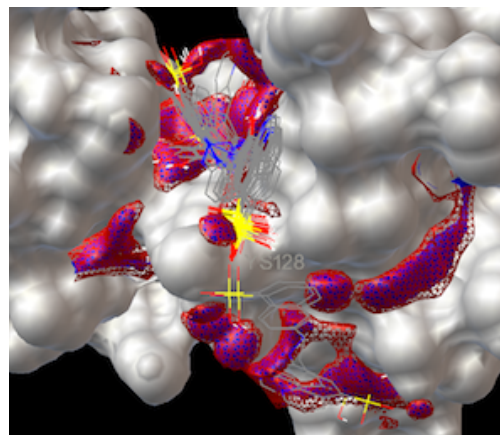


Fig. 91. Red and blue areas are sites that have affinity for nitrogen and oxygen. The same areas have affinity for both nitrogen and oxygen.

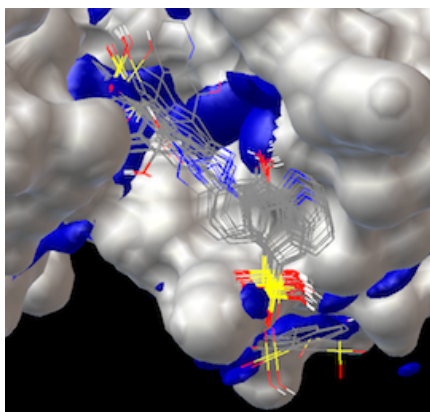


Fig. 92. Nitrogen affinity map.

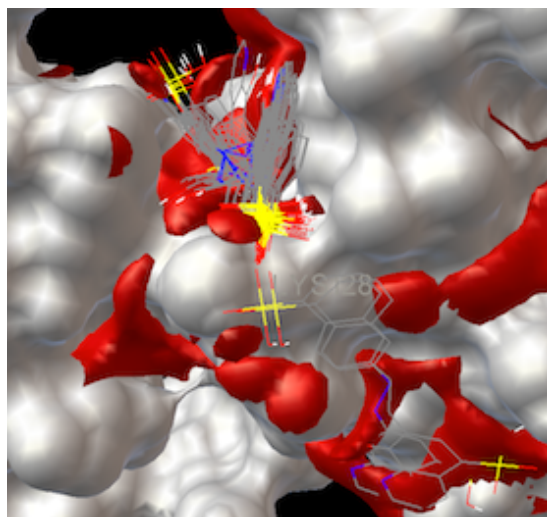


Fig. 93. Oxyg. affinity map. The sulfonic acid groups are buried inside the areas with affinity for oxygen.

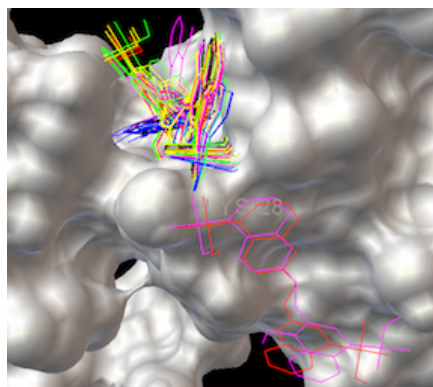


Fig. 94. AB75 – green, AB75-1 – blue, AB75-2 – red, AB75-3 – black, AB75-4 – purple, AB75-5 – yellow.

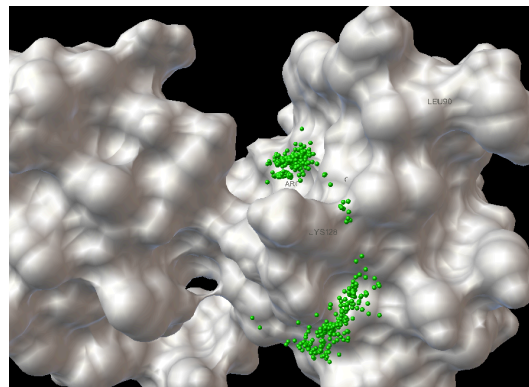


Fig. 95. All 600 conformations represented with green spheres.

Docking of AB75, 75-1 – 75-5 against the R3 region of the DBD in c-Myb

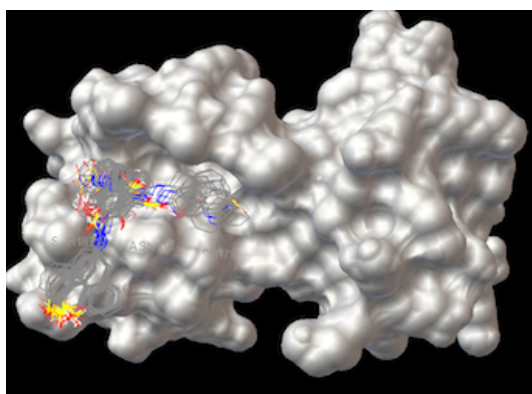


Fig. 96. Top 5 conformations of every compound. One can see that they bind in the same pocket/site.

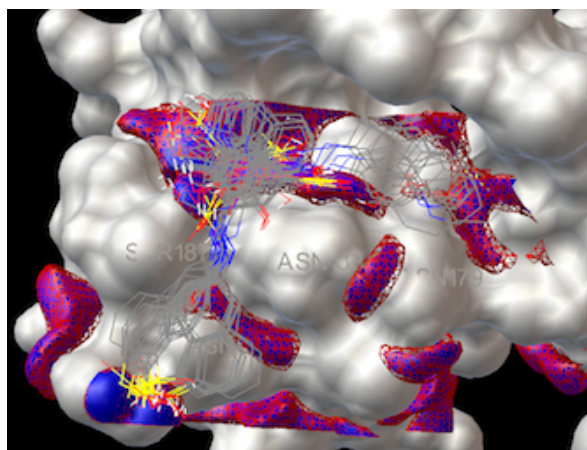


Fig. 97. Red and blue areas are sites that have affinity for nitrogen and oxygen.

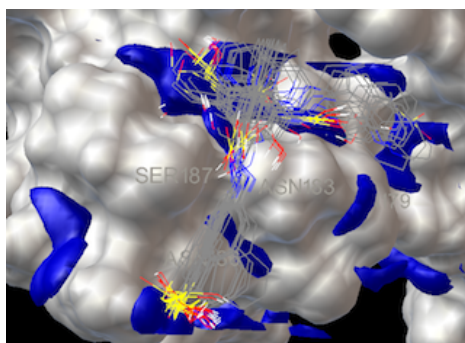


Fig. 98. Nitrogen affinity map.

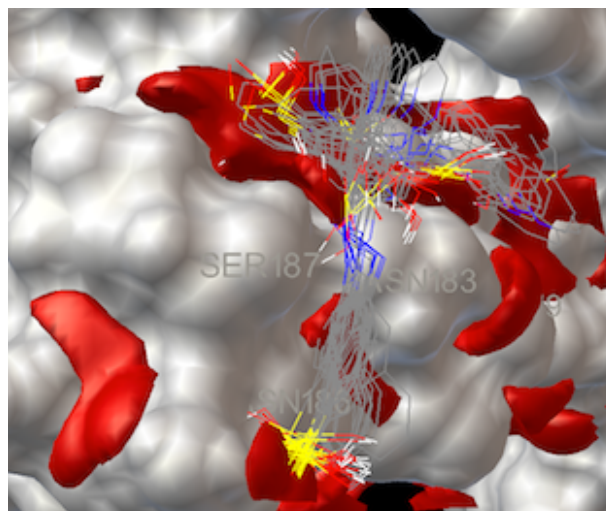


Fig. 99. Oxygen affinity map. The sulfonic acid groups are buried inside the areas with affinity for oxygen.

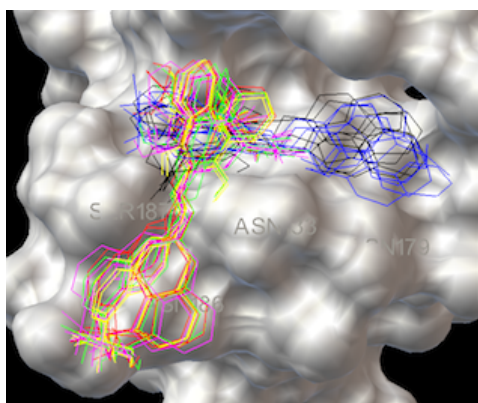


Fig. 100. AB75 – green, AB75-1 – blue, AB75-2 – red, AB75-3 –black, AB75-4 – purple, AB75-5 – yellow.

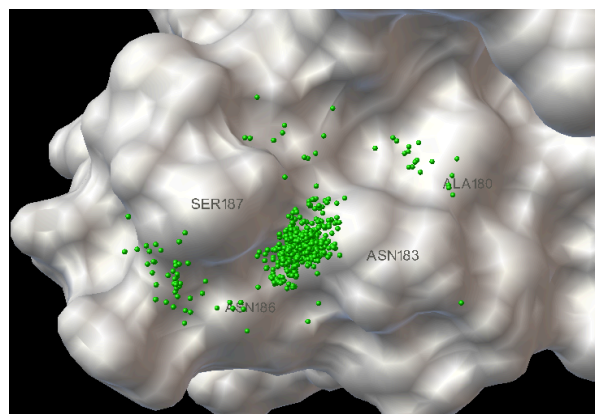


Fig. 101. All 600 conformations represented with green spheres.

Docking of AB76, 76-1 – 76-5 against the R2 region of the DBD in c-Myb

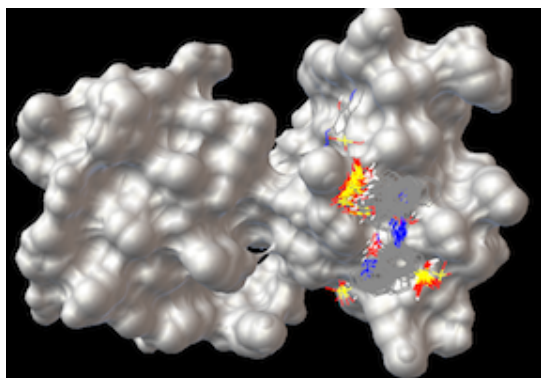


Fig. 102. Top 5 conformations of every compound. One can see that they bind in the same pocket/site.

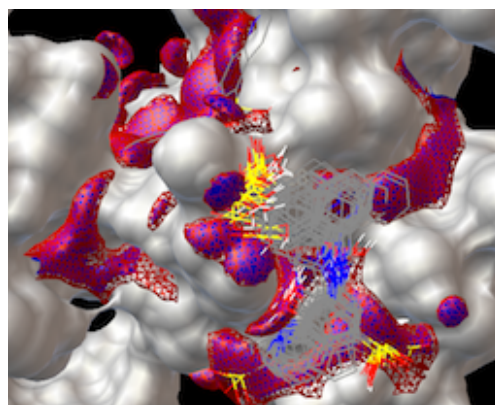


Fig. 103. Red and blue areas are sites that have affinity for nitrogen and oxygen.

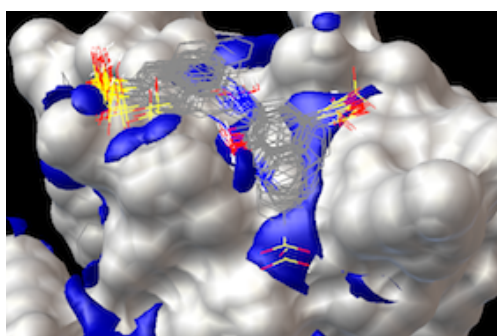


Fig. 104. Nitrogen affinity map.

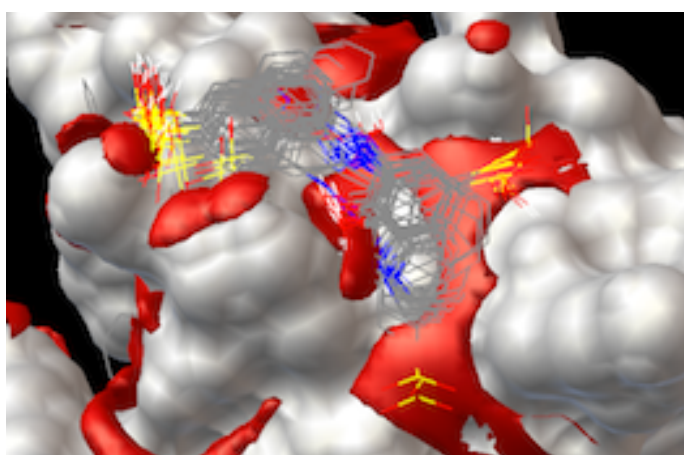


Fig. 105. Oxygen affinity map. The sulfonic acid groups are buried inside the areas with affinity for oxygen.

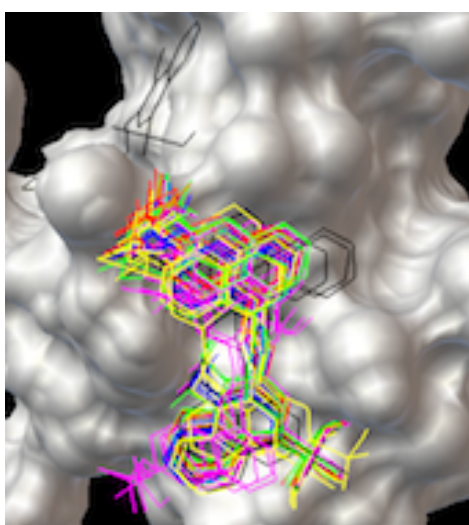


Fig. 106. AB76 – green, AB76-1 – blue, AB76-2 – red, AB76-3 – black, AB76-4 – purple, AB76-5 – yellow.

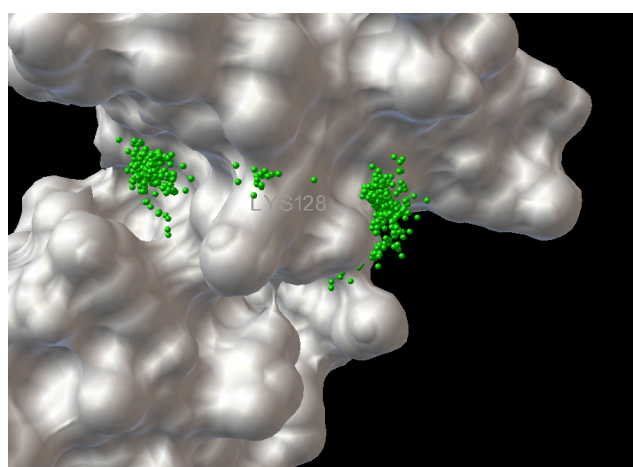


Fig. 107. All 600 conformations represented with green spheres.

Docking of AB76, 76-1 – 76-5 against the R3 region of the DBD in c-Myb

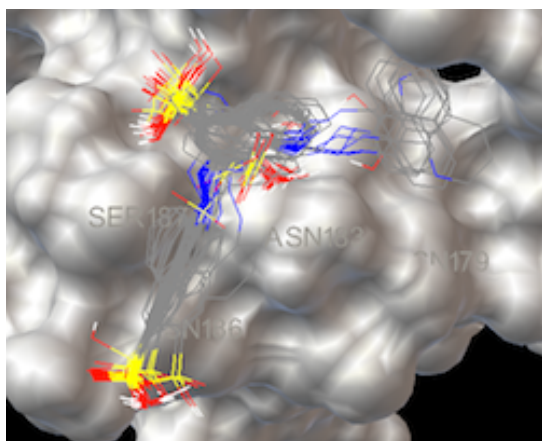


Fig. 108. Top 5 conformations of every compound. One can see that they bind in the same pocket/site.

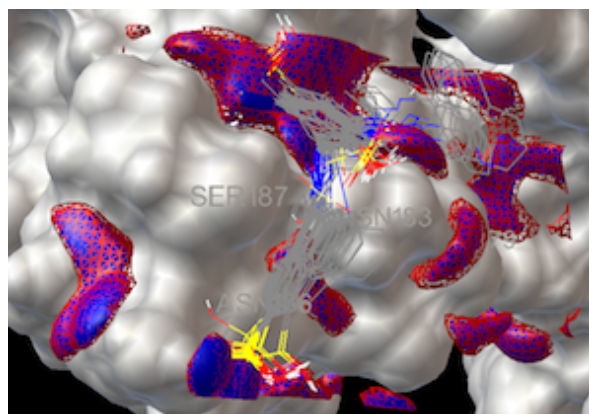


Fig. 109. Red and blue areas are sites that have affinity for nitrogen and oxygen. The same areas have affinity for both atoms.

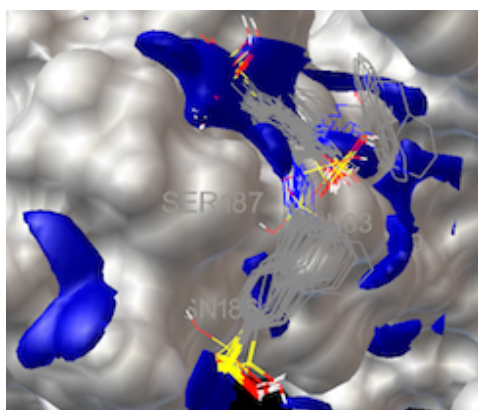


Fig. 110. Nitrogen affinity map.

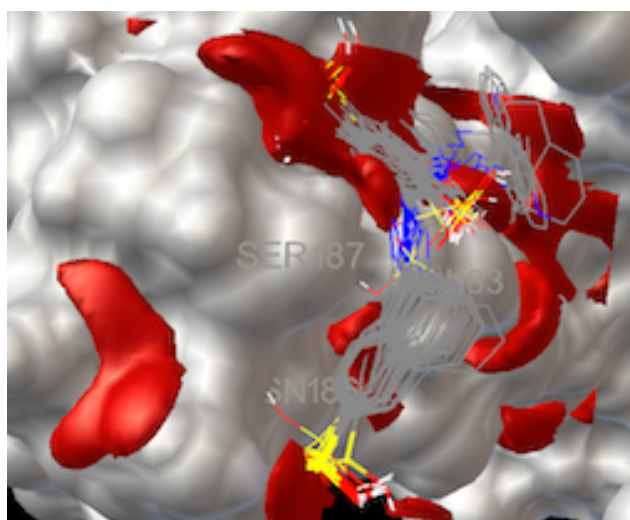


Fig. 111. Oxygen affinity map. The sulfonic acid groups are buried inside the areas with affinity for oxygen.

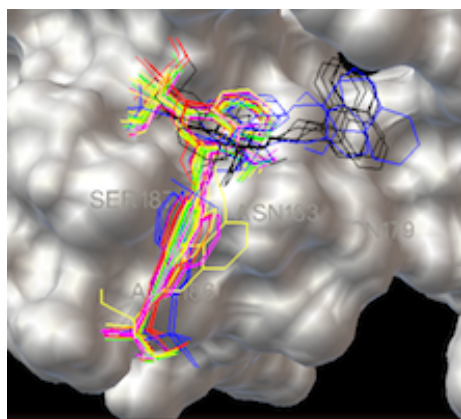


Fig. 112. AB76 – green, AB76-1 – blue, AB76-2 – red, AB76-3 – black, AB76-4 – purple, AB76-5 – yellow.

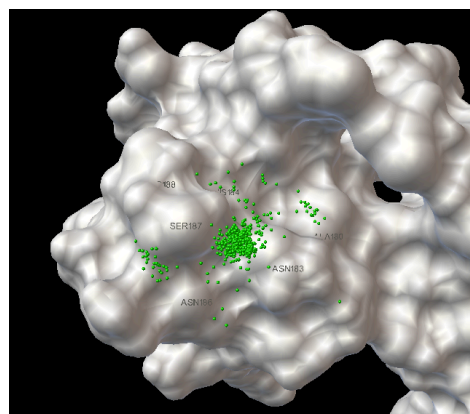


Fig. 113. All 600 conformations represented with green spheres. All conformations are located between Ser187 and Asn183.

Docking of AB77, 77-1 – 77-5 against the R2 region of the DBD in c-Myb

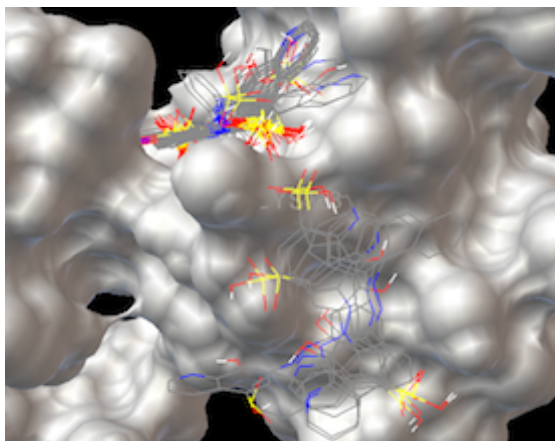


Fig. 114. Top 5 conformations of every compound. One can see that they bind in the same pocket/site.

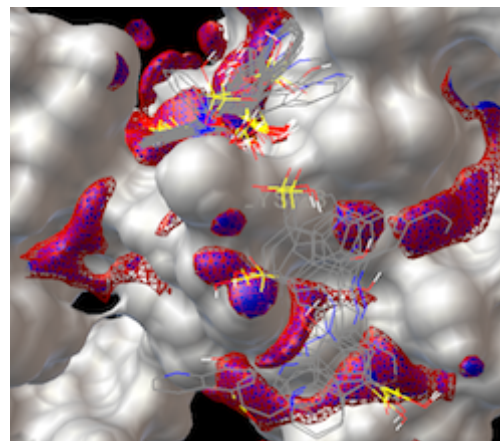


Fig. 115. Red and blue areas are sites that have affinity for nitrogen and oxygen. The same areas have affinity for both atoms.

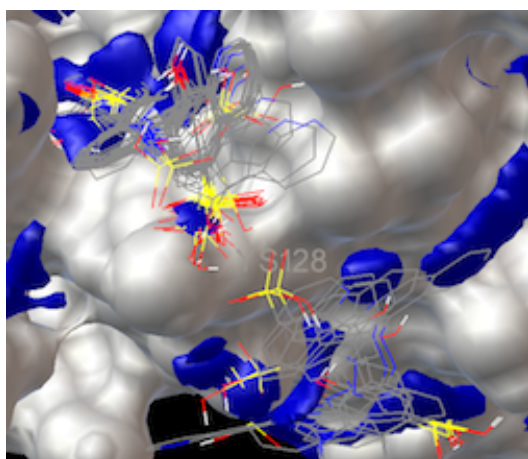


Fig. 116. Nitrogen affinity map.

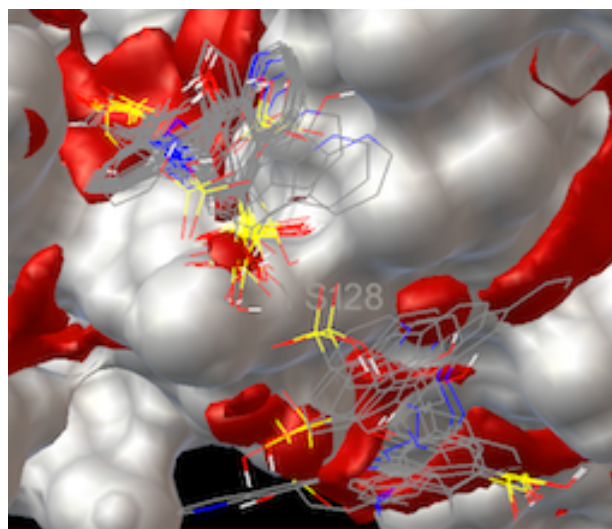


Fig. 117. Oxygen affinity map. The sulfonic acid groups are buried inside the areas with affinity for oxygen.

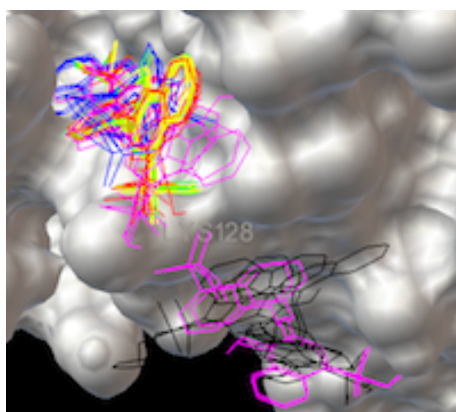


Fig. 118. AB77 – green, AB77-1 – blue, AB77-2 – red, AB77-3 – black, AB77-4 – purple, AB77-5 – yellow.

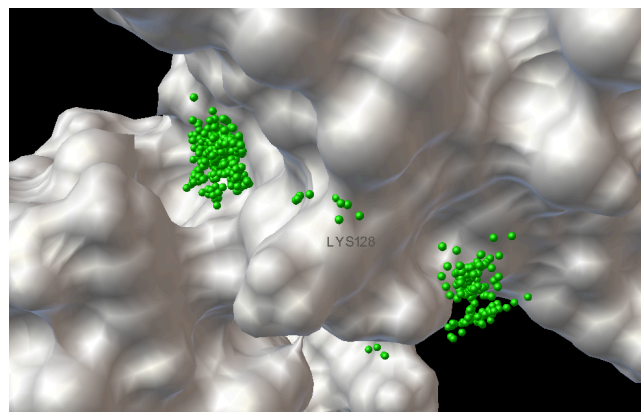


Fig. 119. All 600 conformations represented with green spheres.

Docking of AB77, 77-1 – 77-5 against the R3 region of the DBD in c-Myb

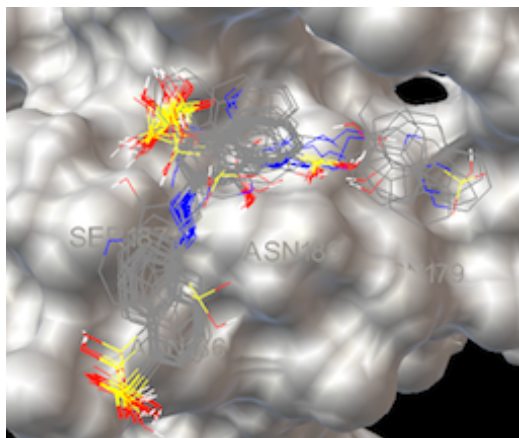


Fig. 120. Top 5 conformations of every compound. One can see that they bind in the same pocket/site.

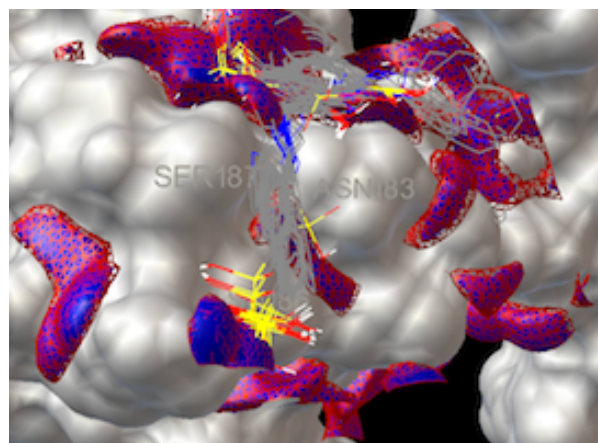


Fig. 121. Red and blue areas are sites that have affinity for nitrogen and oxygen. The same areas have affinity for both atoms.

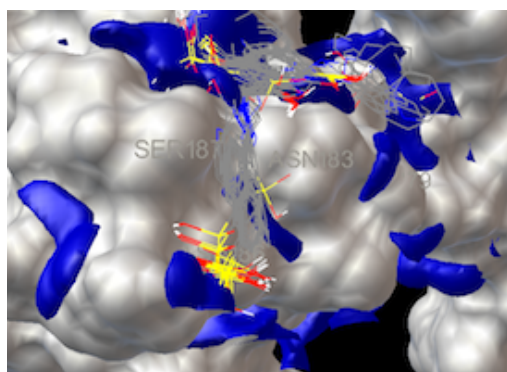


Fig. 122. Nitrogen affinity map.

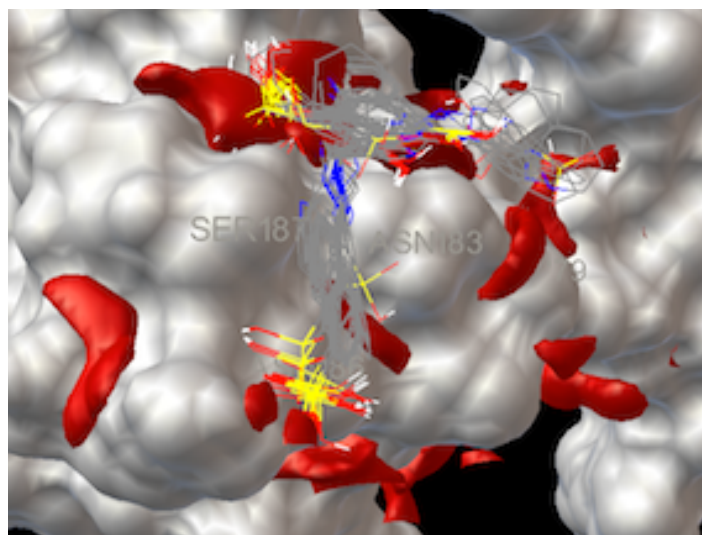


Fig. 123. Oxygen affinity map. The sulfonic acid groups are buried inside the areas with affinity for oxygen.

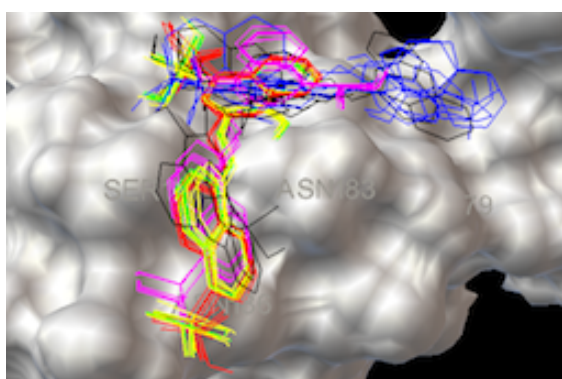


Fig. 124. AB77 – green, AB77-1 – blue, AB77-2 – red, AB77-3 – black, AB77-4 – purple, AB77-5 – yellow.

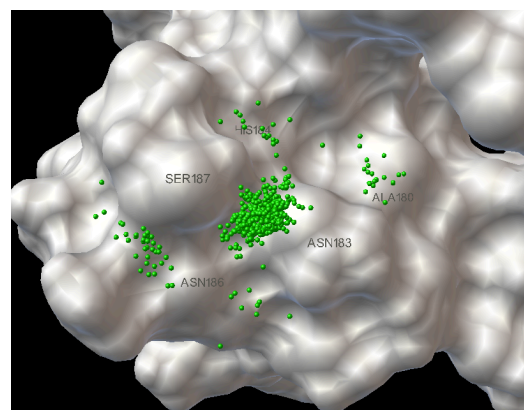


Fig. 125. All 600 conformations represented with green spheres.

Docking of AB78, 78-1 – 78-5 against the R2 region of the DBD in c-Myb

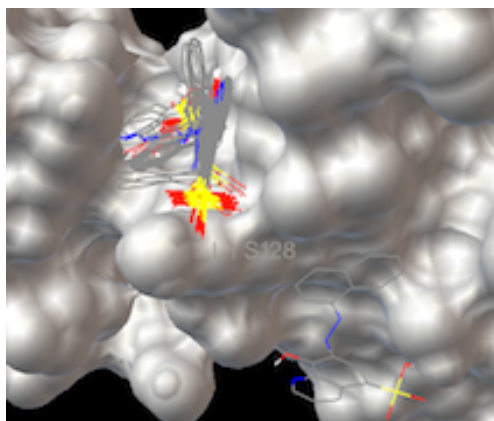


Fig. 126. Top 5 conformations of every compound. One can see that they bind in the same pocket/site.

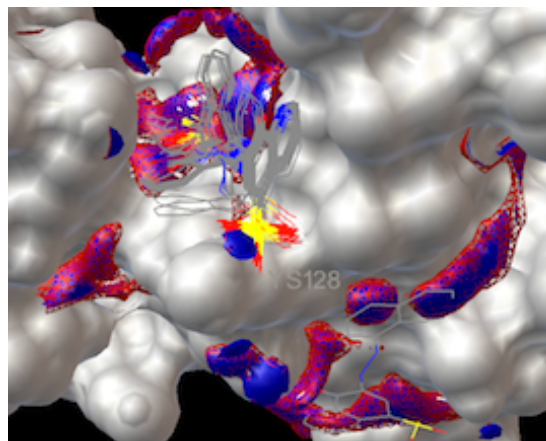


Fig. 127. Red and blue areas are sites that have affinity for nitrogen and oxygen. The same areas have affinity for both atoms.

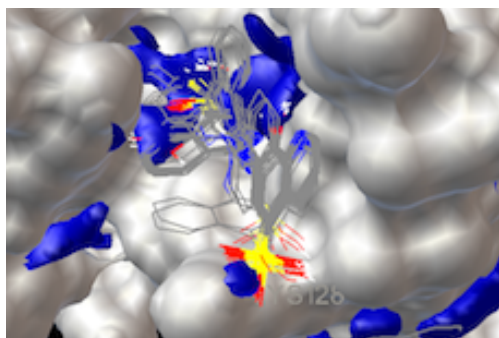


Fig. 128. Nitrogen affinity map.

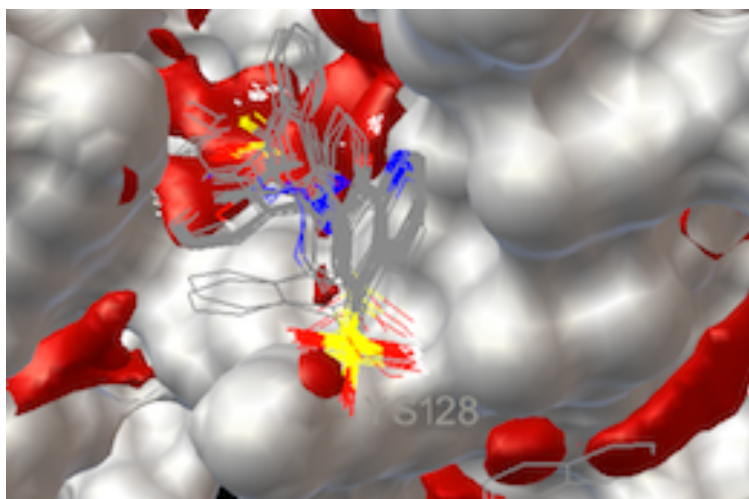


Fig. 129. Oxygen affinity map. The sulfonic acid groups are buried inside the areas with affinity for oxygen.

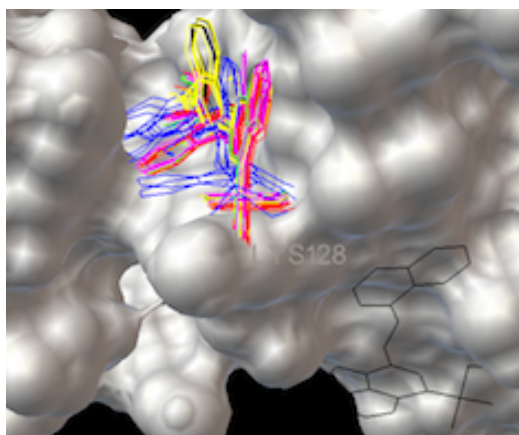


Fig. 130. AB78 – green, AB78-1 – blue, AB78-2 – red, AB78-3 – black, AB78-4 – purple, AB78-5 – yellow.

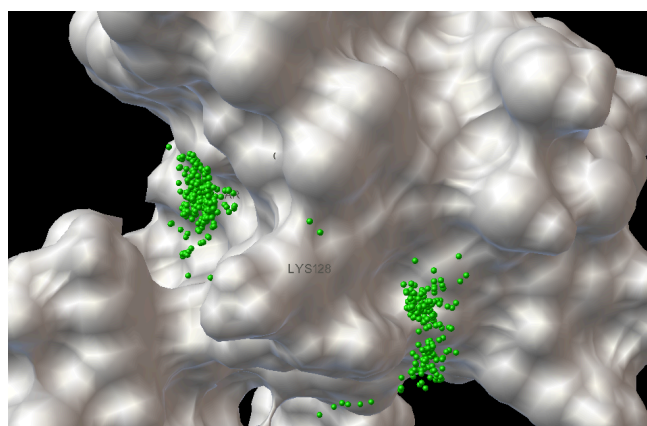


Fig. 131. All 600 conformations represented with green spheres.

Docking of AB78, 78-1 – 78-5 against the R3 region of the DBD in c-Myb

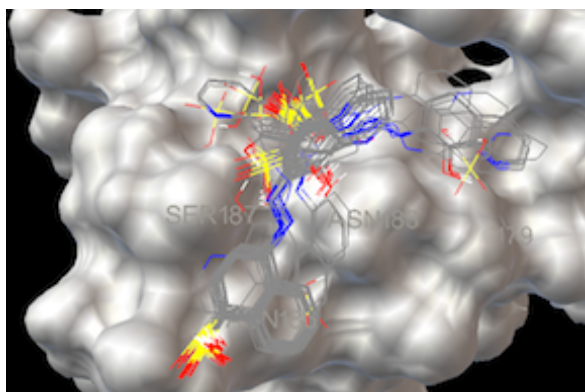


Fig. 132. Top 5 conformations of every compound. One can see that they bind in the same pocket/site.

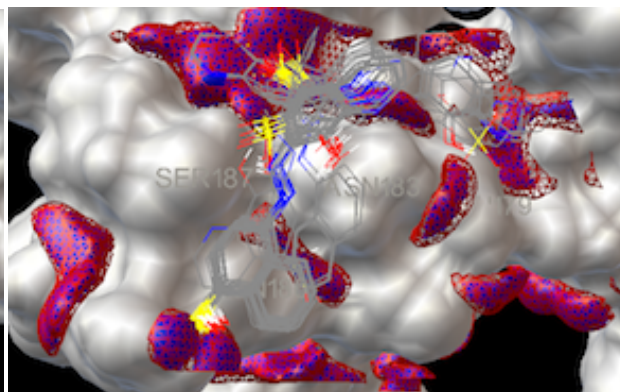


Fig. 133. Red and blue areas are sites that have affinity for nitrogen and oxygen. The same areas have affinity for both atoms.

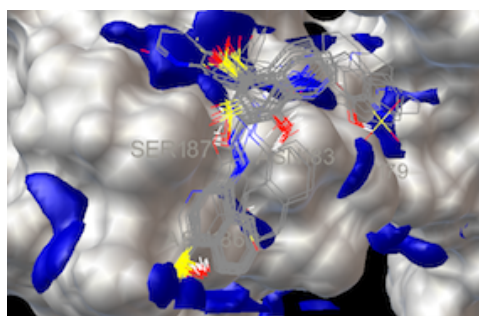


Fig. 134. Nitrogen affinity map.

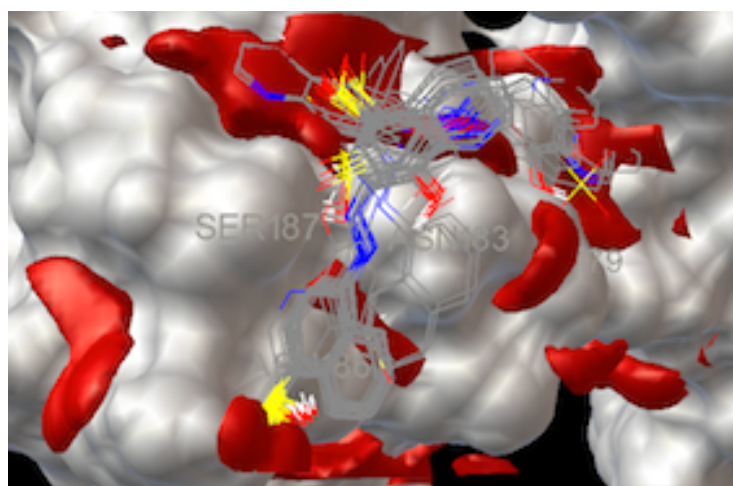


Fig. 135. Oxygen affinity map. The sulfonic acid groups are buried inside the areas with affinity for oxygen.

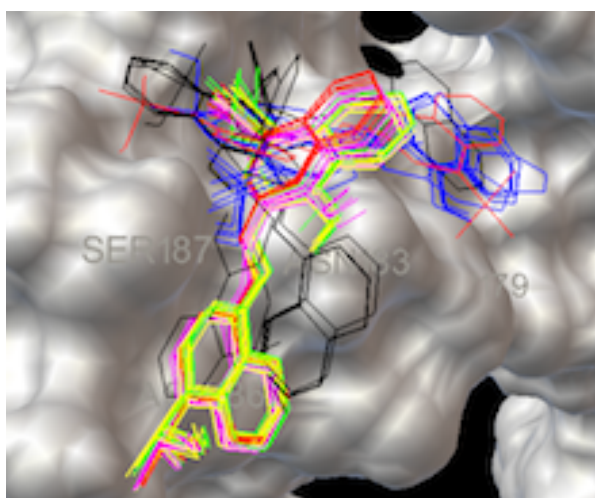


Fig. 136. AB78 – green, AB78-1 – blue, AB78-2 – red, AB78-3 – black, AB78-4 – purple, AB78-5 – yellow.

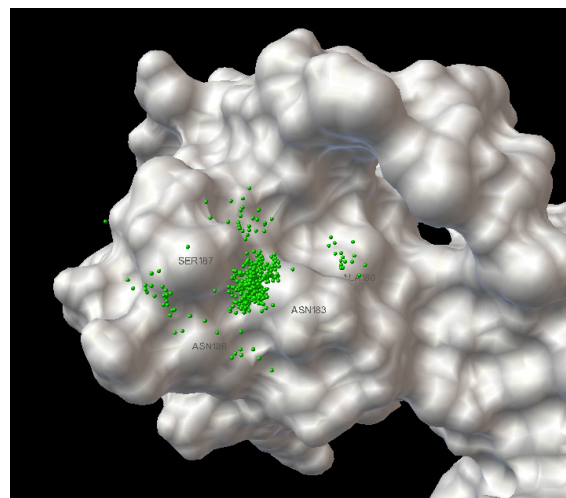


Fig. 137. All 600 conformations represented with green spheres.

Visualization of the re-docked conformations of ABAD45 and ABAD53

Re-docking of ABAD45 in the R2 region in the DBD of the c-Myb

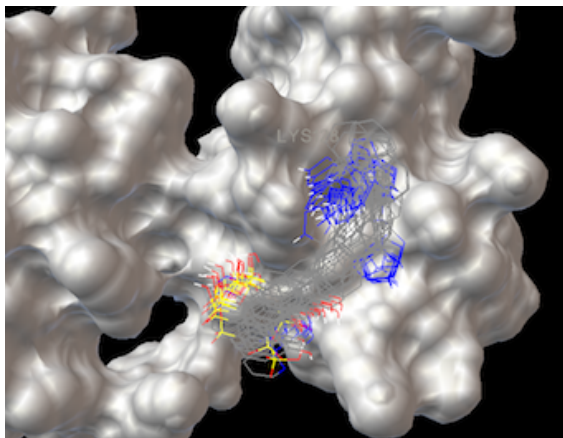


Fig. 138. Top conformations from every single of the 26 runs – 26 conformations totally. All conformations have been docked in the region near Lys128.

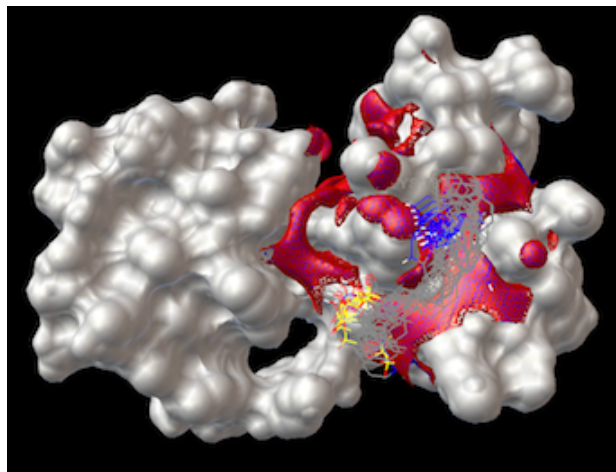


Fig. 139. Red and blue areas are sites that have affinity for nitrogen and oxygen. The same areas have affinity for both nitrogen and oxygen.

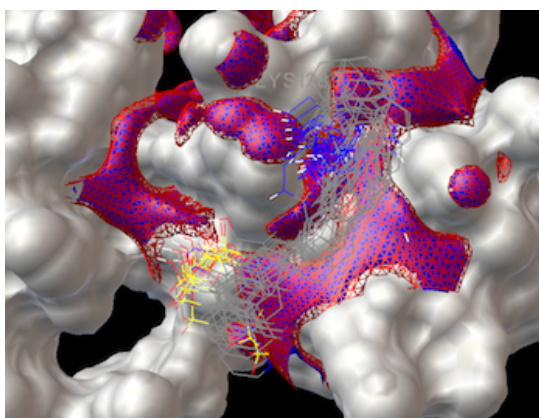


Fig. 140. Same as fig. 139, just a closer view.

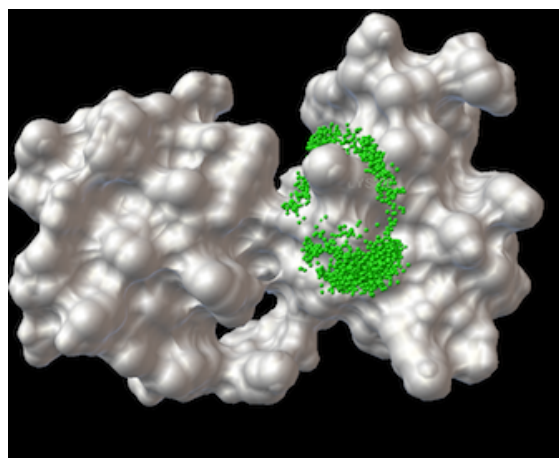


Fig. 141. All 2600 conformations of ABAD45 represented with small green spheres. All the spheres are grouped near Lys128.

Re-docking of ABAD45 in the R3 region in the DBD of the c-Myb

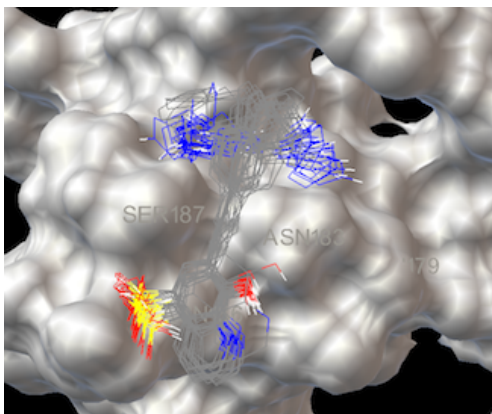


Fig. 142. Top confor. from every single of the 26 runs – 26 confor. totally. All confor. are docked in the same region.

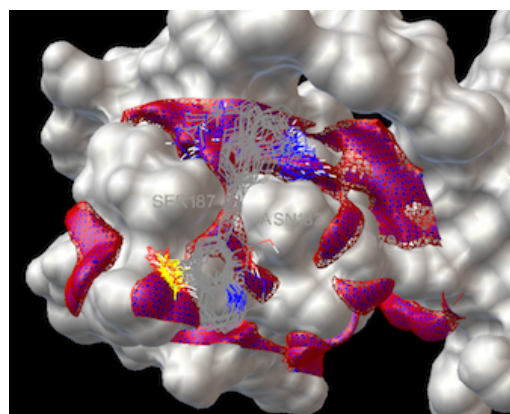


Fig. 143. Red and blue areas are sites that have affinity for nitrogen and oxygen.

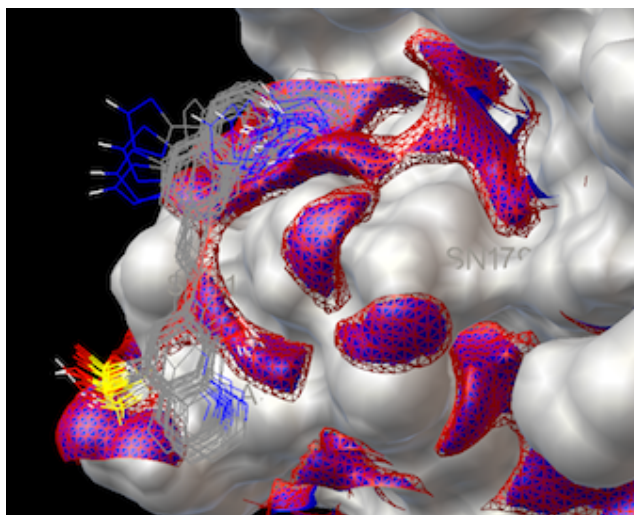


Fig. 144. The same areas have affinity for both nitrogen and oxygen. One can see that the sulfonic acid groups and the tetrazole rings are buried in the regions with affinity for nitrogen and oxygen.

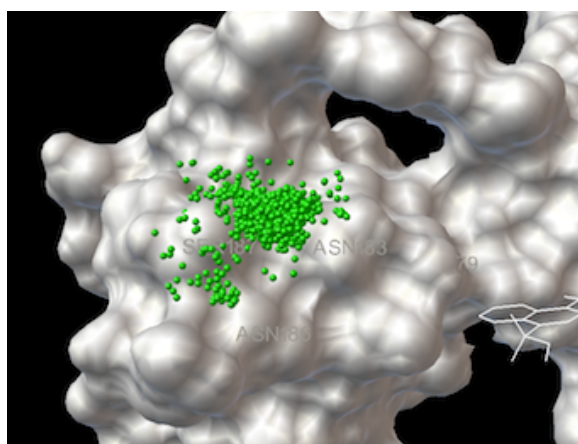


Fig. 145. All 2600 conformations of ABAD45 represented with small green spheres. All the spheres are grouped near Ser187 and Asn183.

Re-docking of ABAD53 in the R2 region in the DBD of the c-Myb

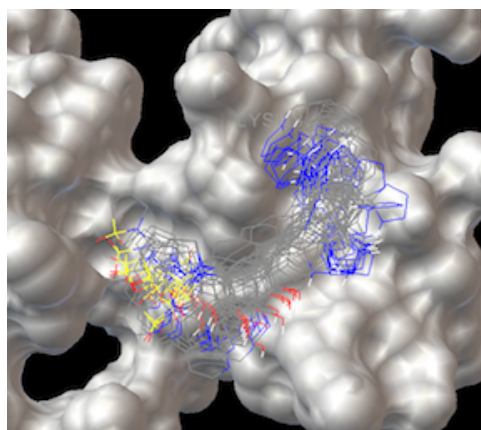


Fig. 146. Top confor. from every single of the 26 runs – 26 confor. totally. All confor. were docked in the same region.

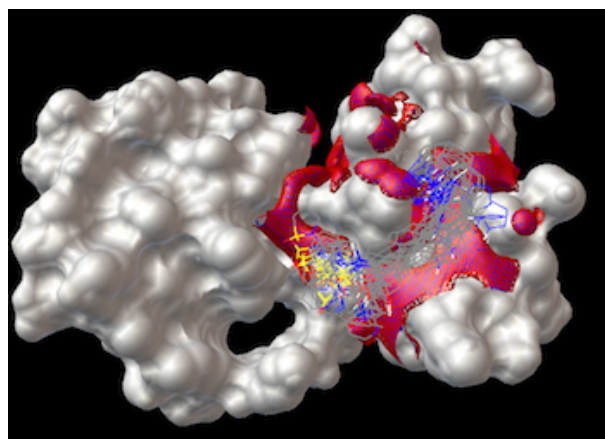


Fig. 147. Red and blue areas are sites that have affinity for nitrogen and oxygen. The same areas have affinity for both atoms.

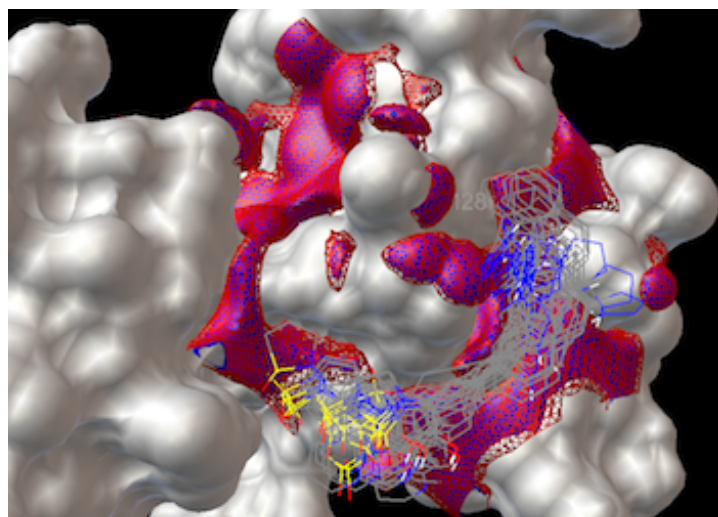


Fig. 148. Red and blue areas are sites that have affinity for nitrogen and oxygen. The same areas have affinity for both nitrogen and oxygen. One can see that oxygen atoms on the sulfonic acid groups and tetrazole rings are buried in the regions with affinity for nitrogen and oxygen.

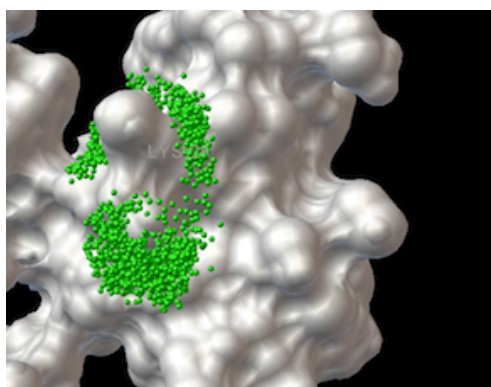


Fig. 149. All 2600 confor. of ABAD53 represented with small green spheres. All the spheres are grouped near Lys128.

Re-docking of ABAD53 in the R3 region in the DBD of the c-Myb

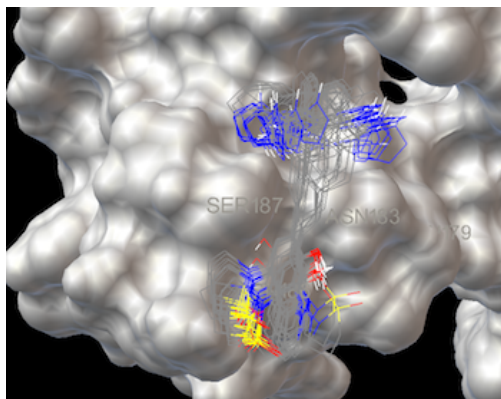


Fig. 150. Top confor. from every single of the 26 runs – 26 confor. totally. All confor. were docked in the same region.

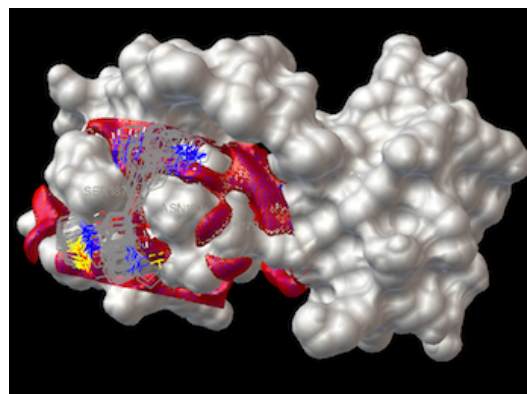


Fig. 151. Red and blue areas are sites with affinity for nitrogen and oxygen. The same areas have affinity for both atoms.

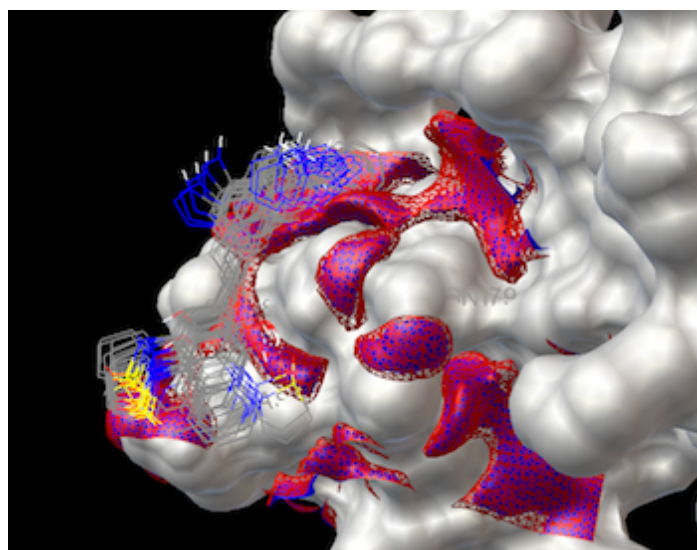


Fig. 152. Red and blue areas are sites that have affinity for nitrogen and oxygen. The same areas have affinity for both nitrogen and oxygen. One can see that oxygen atoms on the sulfonic acid groups and tetrazole rings are buried in the regions with affinity for nitrogen and oxygen.

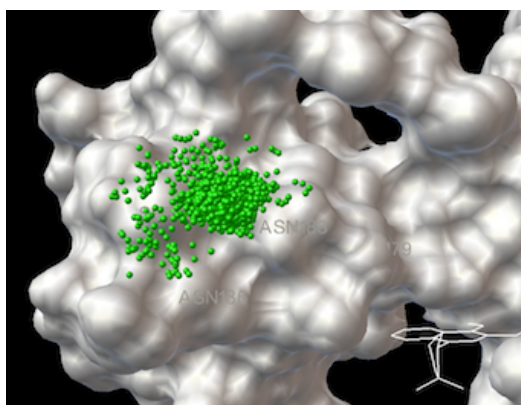


Fig. 153. 2600 conformations of ABAD53 represented with small green spheres that are grouped near Ser187 and Asn183.

References

1. Alberts, B. et al., *Molecular biology of the cell*, 5th ed., Garland Science 2008
2. Reddy, E. P. et al., *Oncogene*, 1999, **18**, p. 3017
3. Ogata, K. et al., *Cell*, 1994, **79**, p. 639
4. Gabrielsen, O. S. et al., *Science*, 1991, **253**, p. 1140
5. Ogata, K. et al., *Nat. Struct. Biol.*, 1995, **2**, p. 309
6. Saether, T. et al., *J. Biol. Chem.*, 2007, **282**, p. 13994
7. Ramsay, R. G. et al., *Nat. Rev., Cancer*, 2008, **8**, p. 523
8. Arndt, H., *Angew. Chem. Int. Ed.*, 2006, **45**, p. 4552
9. Berg, T., *Cur. Opin. in Chem. Biol.*, 2008, **12**, p. 464
10. Schneider, G. et al., *Nat. Rev., Drug Discov.*, 2005, **4**, p. 649
11. Lemke, T. L. et al., *Foey's principles of medicinal chemistry*, 6th ed., Lippincott Williams & Wilkins, 2008
12. a) Morris, G. M. et al., *J. Comput. Chem.*, 1998, **19**, p. 1639
b) Huang, N. et al., *J. Chem. Inf. Model.*, 2006, **46**, p. 243
c) Meng, E. C. et al., *J. Comput. Chem.*, 1992, **13**, p. 505
d) Weiner, P. K. et al., *J. Comput. Chem.*, 1981, **2**, p. 287
e) Nilsson, L., *J. Comput. Chem.*, 1986, **7**, p. 591
f) Brooks, B.R. et al., *J. Comput. Chem.*, 1983, **4**, p. 187
13. a) Wang, R. et al., *J. Mol. Model.*, 1998, **4**, p. 379
b) Krammer, A. et al., *J. Mol. Graph. Model.*, 2005, **23**, p. 395
c) Wang, R. et al., *J. Comput. Aided Mol. Des.*, 2002, **16**, p. 11
14. a) Muegge, I., *J. Med. Chem.*, 2006, **49**, p. 5895
b) Gohlke, H. et al., *J. Mol. Biol.*, 2000, **295**, p. 337
c) Velec, H. F. G. et al., *J. Med. Chem.*, 2005, **48**, p. 6296
d) Ishchenko, A.V. et al., *J. Med. Chem.*, 2002, **45**, p. 2770
e) Mitchell, J. B. O. et al., *J. Comput. Chem.*, 1999, **20**, p. 1165
f) Mooij, W. T. et al., *Proteins*, 2005, **61**, p. 272
g) Yang, C. Y. et al., *J. Med. Chem.*, 2006, **49**, p. 5903
15. a) Garrett, M. M. et al., *Curr. Prot. in Bioinf.*, 2008, **unit 8.14**, p. 1
b) Cosconati, A. et al., *Expert Opin. Drug Discov.*, 2010, **5**, p. 597
16. Huey, R. et al., *J. Comput. Chem.*, 2007, **28**, p. 1145

17. Garrett, M. M. et al., *J. Comput. Chem.*, 1998, **19**, p. 1639
18. www.esilrch1.esi.umontreal.ca/~bioinfo/procedures/LOGICIELS/MANUELS/AutoDock3.0.5_UserGuide.ps.
19. Chang, M. W. et al., *J. Comput. Chem.*, 2008, **29**, p. 1753
20. Rosenfeld, R. J. et al., *J. Comput. Aided. Mol. Des.*, 2003, **17**, p. 525
21. www.ucl.ac.uk/~rmhasek/misc/autodock.pdf
22. Hopkins, A. L. et al., *Drug Discov. Today*, 2004, **9**, p. 430
23. Li, C. et al., *J. Med. Chem.*, 2004, **47**, p. 6681
24. www.bmb.uga.edu/bcmb8110/ADT_Autodock_Tutorial.pdf
<http://amit1b.files.wordpress.com/2008/04/autodock4.pdf>
http://rwjms1.umdj.edu/kholodvl/tut/ADT_guide.pdf
<http://chemistry.umeche.maine.edu/Modeling/AutoDock-tutorial.pdf>
25. James, S. F. et al., *Anal. Chem.*, 1961, **33**, p. 1381
26. James, S. F. et al., *Anal. Chem.*, 1957, **29**, p. 821
27. Gangadhar, C., *Chemosphere*, 1996, **32**, p. 267
28. Lipinski, C. A. et al., *Adv. Drug. Deliv. Rev.*, 2001, **46** p. 3
29. a) Navia, M. A. et al., *Drug Discov. Tod.*, 1996, **1**, p. 179
b) Pardridge, W. M., *Adv. Drug Deliv. Rev.*, 1995, **15**, p. 5
30. Chang, M. W. et al., *J. Chem. Inform. Model.*, 2007, **47**, p. 1258
31. a) Langmuir, I., *J. Am. Chem. Soc.*, 1919, **41**, p. 868
b) Langmuir, I., *J. Am. Chem. Soc.*, 1919, **41**, p. 1543
32. Burger, A., *Prog. Drug. Res.*, 1991, **37**, p. 287
33. Patani, G. A. et al., *Chem. Rev.*, 1996, **96**, p. 3147
34. a) Thornber, C. W., *Chem. Soc. Rev.*, 1979, **8**, p. 563
b) Singh, H. et al., *Prog. Med. Chem.*, 1980, **17**, p. 151
c) Butler, R. N., *Adv. Het. Chem.*, 1977, **21**, p. 323.
d) Wittenberger, S. J., *Org. Prep. Proced. Int.*, 1994, **26**, p. 499
35. Butler, R. N. et al., *Compreh. Het. Chem. II*, vol. **4**., Eds. Pergamon Oxford, 1996
36. a) Wexler, R. R. et al., *J. Med. Chem.*, 1996, **39**, p. 625
b) Noda, K. et al., *J. Biol. Chem.*, 1995, **270**, p. 2284
c) Kubo, K. et al., *J. Med. Chem.*, 1993, **36**, p. 2182
37. a) Nohara, A. et al., *J. Med. Chem.*, 1979, **22**, p. 290
b) Chando, T. J. et al., *Drug Metab. Dispos.*, 1998, **26**, p. 408
c) Perrier, L. et al., *J. Pharmacol. Exp. Ther.*, 1994, **271**, p. 91
38. Narayan, V. A. et al., *J. Bio. Chem.*, 1997, **272**, p. 7801
39. Shinichiro, O., *Biochem.*, 2004, **43**, p. 16027
40. Zollinger, H., *Diazo Chemistry*, VCH, 1994

41. Zollinger, H., *Color Chemistry*, VCH, 1987
42. Kosynkin, D. et al., *J. Chem. Soc., Perkin Trans. 2*, 1997, p. 2003
43. Zollinger, H. et al., *Azo Coupling Reactions: Structures and Mechanisms*, Technisch-Chemisches Laboratorium, ETH
44. Falomir, E. et al., *Tetrahedron Lett.*, 2003, **44**, p. 539
45. a) Stephens, R. D. et al., *J. Org. Chem.*, 1963, **28**, p. 2163
 b) Stephens, R. D. et al., *J. Org. Chem.*, 1963, **28**, p. 3313
46. Cassar, L., *J. Organomet. Chem.*, 1975, **93**, p. 253
47. Dieck, H. A. et al., *J. Organomet. Chem.*, 1975, **93**, p. 259
48. Sonogashira, K. et al., *Tetrahedron Lett.*, 1975, **16**, p. 4467.
49. a) Chinchilla, R. et al., *Chem. Rev.*, 2007, **107**, p. 874
 b) Organ, M. G. et al., *J. Org. Chem.*, 2004, **69**, p. 695
 c) Huang, Q. et al., *J. Org. Chem.*, 2002, **67**, p. 3437
50. (a) Sonogashira, K. et al., *Tetrahedron Lett.*, 1975, **16**, p. 4467
 (b) Sonogashira K., *Handbook of Organopalladium Chem. for Org. Synth.*, Wiley: New York, 2002
 (c) Sonogashira, K., *J. Organomet. Chem.*, 2002, **653**, p. 46
 (d) Tykwinski, R. R., *Angew. Chem. Int. Ed.*, 2003, **42**, p. 1566
51. a) Amatore, C. et al., *Organometallics*, 1993, **12**, p. 3168
 b) Grosshenny, V. et al., *J. Org. Chem.*, 1997, **62**, p. 1491
52. a) Singh, R. et al., *J. Org. Chem.*, 1989, **54**, p. 4453
 b) Austin, W. B. et al., *J. Org. Chem.*, 1981, **46**, p. 2280
53. a) Alami, M. et al., *Tetrahed. Lett.*, 1993, **34**, p. 6403
 b) Houpis, I. N. et al., *Tetrahed. Lett.*, 1994, **35**, p. 9355
 c) Boydston, A. J., *J. Org. Chem.*, 2002, **67**, p. 8812
54. a) Larock, R. C., *Comprehensive Organic Transformations*, VCH, New York, 1989
 b) Sandler, S. R. et al., *Organic Functional Group Preparation*, Academic Press, New York, 1968
 c) Kabalka, G. W. et al., *Compreh. Org. Synth.*, **Vol. 8**, Pergamon Press, Oxford, 1991
 d) Hudlicky, M., *Reductions in Org. Chem.*, 6th ed., American Chemical Society, Washington DC, 1996
 e) Rylander, P., *Catalytic Hydrogenation in Organic Synthesis*, Academic Press, New York, 1979
 f) Fox, B. A. et al., *Org. Synth.*, 1973, **5**, p. 346
 g) Dirmoth, K. et al., *Org. Synth.*, 1973 **5**, p. 1130
 h) Ickk, R. N. et al., *Org. Synth.*, 1955, **3**, p. 59
 i) Allene, G. F. H. et al., *Org. Synth.*, 1955, **3**, p. 63
55. Li, J. et al., *J. Org. Chem.*, 2010, **75**, p. 2966
56. Wei, L. et al, *Tetrahed. Lett.*, 2003, **44**, p. 1979
57. Brown, C. H. et al., *J. Org. Chem.*, 1986, **51**, p. 4512

58. a) Nakamura, H. et al., *Chem. Med. Chem.*, 2006, **1**, p. 729
b) Lawrence, N. J. et al., *Synth.*, 1999, p. 1656
59. a) Giraud, A. et al., *Tetrahed. Lett.*, 2008, **49**, p. 1107
b) Hamze, A. et al., *Org. Lett.*, 2005, **7**, 2005, p. 5625
c) Hamze, et al., *Synth.*, 2007, **13**, p. 2025
60. Marcienec, B. et al., *Hydrosilylation: A Comprehensive review on Recent Advances*, Springer, 2009
61. Moriarty, R. M. et al., *J. Am. Chem. Soc.*, 1975, **15**, p. 5603
62. Trost, B. M. et al., *J. Am. Chem. Soc.*, 2002, **124**, p. 7922
63. a) Lindlar, H., *Helv. Chim. Acta.*, 1952, **35**, p. 446
b) Lindlar, H. et al., *Org. Synth. Coll.*, 1973, **5**, p. 880
64. Curran, T. T. et al., *Tetrahed. Lett.*, 1995, **36**, p. 4761
65. Garcia-Mota, M. et al., *Theor. Chem. Acc.*, 2010, **128**, p. 663
66. Trost, B. et al., *Compreh. Org. Synth.*, **vol. 8**, Pergamon Press, 1991
67. Nakorn, N. et al., *Org. Lett.*, 2009, **11**, p. 2768
68. Martinez, M. C. R., "Identification of small molecule inhibitors of the transcription factor c-Myb through structure-based drug design and EMSA analysis". Matematisk – naturvitenskaplig fakultet, IMBV, Oslo, UiO, 2009.
69. Mugass, G., "Potensielle inhibitorer av transkripsjonsfaktoren c-Myb analysert *in vitro* og *in vivo*". Matematisk – naturvitenskaplig fakultet, IMBV, Oslo, UiO, 2011.
70. TOCRIS Bioscience, private communication
71. a) Corey, E. J. et al., *Tetrahed. Lett.*, 1972, **13**, p. 3769
b) Desai, N. B. et al., *J. Am. Chem. Soc.*, 1962, **84**, p. 1745
72. Yan, J. et al., *Bioorg. Med. Chem. Lett.*, 2002, **12**, p. 2209
73. Nelson, D. L. et al., *Lehninger Principles of Biochemistry*, 4th ed., WH Freeman and Company, 2005.
74. Cross, J. B. et al., *J. Chem. Inf. Model.*, 2009, **49**, p. 1455
75. Pavia, L. D. et al., *Introduction to Spectroscopy*, 4th ed., Brooks/Cole, Cengage Learning, 2009.
76. a) Wang, Y. et al., *J. Chem. Res.*, 2007, **12**, p. 728
b) Yeom, C. et al., *Synlett*, 2008, **4**, p. 565
c) Sakata, T. et al., *Langmuir*, 2007, **23**, p. 2269
d) Bell, M. G. et al., PCT Int. Appl., 2007140174, Dec., 2007
e) Yokoyama, A. et al., *Org. Lett.*, 2008, **10**, p. 3207