Tesi Doctoral

NMR IN DRUG DISCOVERY. FROM SCREENING TO STRUCTURE-BASED DESIGN OF ANTITUMORAL AGENTS

Ricard A. Rodríguez Mias

Departament de Química Orgànica

Facultat de Química

Universitat de Barcelona

Barcelona, Juliol 2006

Memòria presentada per:

Ricard A. Rodríguez Mias

Per optar al títol de Doctor per la Universitat de Barcelona

Revisada per:

Dr. Ernest Giralt Lledó

Universitat de Barcelona

Director

Programa de Química Fonamental-Química Orgànica

Bienni 2000-2002

Departament de Química Orgànica

Facultat de Química

Universitat de Barcelona

Barcelona, Juliol 2006

Als meus pares

i a la meva germana.

Nothing is impossible. Not if you can imagine it.

That's what being a scientist is all about.

Prof. Hubert Farnsworth

No, that's what being a magical elf is all about.

Cubert J. Farnsworth

Agraïments

AAAAAAAAAAAAAAAAAARRRRRRRRRRGGGGGGGGGHHHHHHHHHHHHHHHHHHHHHHHHHHHH.!!!!!!!!!!!!!!!!!

Bueno, acabeu de presenciar una de les últimes contraccions d'aquest part, per tant suposo que ha arribat el moment dels agraïments; de fet ja que hi som m'agradaria presentar una queixa per que l'epidural que m'han posat no ha acabat de fer efecte.

En primer lloc m'agradaria agrair a l'Ernest el fet d'haver compartit la paternitat d'aquesta criatura amb mi, gràcies pel teu suport en aquests anys així com per la llibertat que m'has donat per portar a terme totes les idees boges, submarinos i col·laboracions. Al Fernando, tot i que no figura com a un dels pares d'aquesta tesi, m'agradaria agrair-te el fantàstic paper de comadrona durant aquests últims mesos, però també els consells, els ànims i els cafès de les tardes de cap de setmana al laboratori.

I'm also very grateful to Dr Pellecchia and Dr. Wagner for their kindness and hospitality during my visits to their respective laboratories. Both stays were scientifically very enriching, but also very fun and are probably the reason of my Peter Pan syndrome and my post doc vocation. For this I should also thank the people in both laboratories: Mari Lisa, Sina and Jamie in San Diego and Tom, Hari, Federico, Mallika, Monica, Alison and Michael… in Boston.

Si fem una mica de memòria històrica, ara que es porta tant … . Tot va començar fa uns anys a Girona, jo era un xavalin pueblerino -ho continuo sent- recent sortit de la Facultat de Ciències dela UdG. En aquell temps ja era un indecís i no tenia clar cap on fer anar la meva carrera i l'Eduard Bardaji amb va acollir durant un temps i introduir al món dels pèptids -gràcies-. El cas es que volia deixar de ser pueblerino i marxar, però com a bon cagat que sóc, no vaig anar gaire lluny: Barcelona. I així es com vaig anar a parar a una comuna Hippy que es feia dir Grup 10. L'Ernest em va seure al costat dels HPLCs entre la gent amb mes classe del grup: els RMNeros. Gràcies al Pau, a l'Anna i al Xavier vaig aprendre a concentrarme entre el soroll de shimatzus infernals, a barallar-me amb la RMN de pèptids, a expressar proteïnes i a gestionar les necessitats d'ordinador i d'impressora d'una marabunta de gent.

En aquella comuna hippy la distracció era constant, i es dividia per igual entre el grup d'oligos: la Berta, el Joan, la Cris Ausin, el Toni, el Dani …, la gent del servei de síntesi de pèptids: Mari Luz, Alberto Adeba, Mele …, el caos del Lab 1, el Lab 5 de la incomparable Lore i l'únic reducte de serenitat, el Lab3 amb la Lia com a gran mama; i com no la Marta Molina intentant posar ordre al caos. Gràcies a tots pels bons moments a la facultat … mai he entès com amb tan poc espai hi podia haver tan bon rollo.

Finalment l'esperada mudança al Parc Científic va arribar, allà encara érem més bons, teníem més classe, més equips i més espai; tot i això la diversió va continuar. El trasllat al Lab300 va ser una aventura interessant, i amb en Xavier vaig aprendre a disfrutar dels virus, les impressores rebels i de l'ajuda inestimable dels amics del SIC. El laboratori encara conservava cert hippisme del Grup10 i gent com la Montse-Montse, la Jimena, la Laia Solé i el seu PIMCO, la Gloria Sancli contagiosament hiperactivitat, el llegendari Kalamar amb qui tant he rigut … tots ells em van ajudar a fer la vida molt fàcil.

I si abans de la mudança els HPLCs formaven part de la meva vida; el destí, com si volgués recordar-me els meus orígens, em va tornar a posar en un altre passadís amb vistes, aquest cop però les vistes eren Waters™. Com a guerrillers en aquest creuament de civilitzacions tenen menció d'honor: el Nacho, un crack en les seves facetes de Dr Jekill i Mr Hide, un acha de l'evolució, del duro i el millor adminsitrador del caos informàtic que hem tingut des d'en Xavier. La Pilarin, pianista, madridista i del PSC, però malgrat tot la millor companya de penes en el càlcul d'estructures de pèptids. El Marcelo, señor del caos, de los mil proyectos y colaboradores, que grande eres. I como no les incorporacions mes recents: el Nessim i la Steph.

I tot i no haver viscut al passadís cal que recordi algunes de les velles glories del grup: el piratilla Pastor, la senyora de les fibres Lolita, la S. Frutos, hazote de doctorandos novatos, companya de batalles contra la seguretat del Parc i font

inesgotable de galetes; la Txell, la jefecilla de latigo mes llarg, el Marc Martinell amb qui vaig compartir l'Honda Accord a San Diego … tots ells han sigut en gran part responsables que em costés tant marxar del laboratori.

Tot seguit es van anar afegint a la festa nous fitxatges, i tot i no estar curtits per les vivències de la facultat, l'Ester, Silvia P, Susana, Eduard, Muriel, Laia, Giovana, Birgit, Roger, Mortezaa han aconseguit que se segueixi respirant un un ambient de laboratori fantàstic. Com no, també m'agradaria recordar la Natalia i el Sergio, la gent del despatxet (la cookie room) per la seva paciència amb els meus saquejos a la cerca de xocolata i sucre. A la Tere, per deixar-se robar l'assortit de plantes xineses, així com pels seus consells de biologia molecular; a la Eva per evitar que el laboratori sigui un caos, i que la paperassa ens inundi. También me gustaría mencionar al digno heredero del proyecto VEGF: Peter, darte las gracias por tu vitalidad, por tu punto de locura y por intentar que mis esfuerzos en el laboratorio sirvan de algo.

Com a conseqüéncia de la situació estratègica del passadís RMNero he pogut disfrutar igualment de de la fantàstica fauna dels dominis del Fernando: l'incomparable humor del Jan, la quasi exasperant curiositat del Tommaso, les ganes de festa d'en Javi payopony, el Jesús "alias sexus¨, i la sempre somrient Martina. La llista es quasi interminable, sobretot si considerem l'area d'influència que inclou el Lab100 amb l'accent andalús del Dani, Pharmamar (anteriorment conegut com el harén de Luis Javier) i Lylliput; gràcies a tots. Com no, també la gent del laboratori de UQSex que ha sigut una gran font de diversió i colaboracions; gràcies Miriam pels teus consells i els ànims. Gràcies també al Paquito, tot i que de vegades m'intenti colar el mateix compost dos cops i al Marc V per la seva paciència i convicció en les meves idees de bombero.

Gran part de la feina d'aquesta tesi però no ha passat a cap d'aquests laboratoris, mes aviat als domins del Servei de RMN. Ells també han patit el meu caos, i han sigut víctimes de moltes de les meves idees boges. A tota la gent del servei els hi vull donar les gràcies per la seva ajuda amb els aparells, per perdonar el desordre que sovint queda darrera meu als espectròmetres i per deixar que m'apropiés dels vostres dominis en les "hora dels manolos". En especial gràcies a la Marga per estar sempre disponible i per escoltar les meves idees, encara que siguin impracticables. També he okupat el laboratori del Dr. Pons, i la seva gent m'ha acollit sempre amb hospitalitat en les excursions a la vitrina de flux laminar, a la cambra freda i a la sala d'incubadors. Gràcies a tots: Jesús, Eric, David, Jascha, Tiago, Pau, Oscar.

Als companys de dinars, Jose P, Albert, Unai, Jesús … els hi vull agrair els bons moments passats mentre travessàvem el desert, les discussions sobre política internacional i els cafès amb les millors vistes de la UB.

I com que en aquests anys hi ha hagut temps per tot, m'agradaria recordar a tots els que m'heu acompanyat en les activitats extraescolars nocturnes, gràcies per les festes dels divendres. Gràcies a les "germanes" Fayna i Mireia – l'esperit més jove del lab - per compartir el vostre capital social i acompanyar-me en la meva època de "Rodriguez"; al Javi payopo per compartir el seu fetge i fer-nos d'amfitrio als San Fermines. Gràcies també a la Luengo, l'Ari, la Natalia, l'Ester, al Javi ra ra ra, … a tots ells per ajudarme a conéixer Barcelona, almenys de nit.

També hi ha hagut temps de participar en activitats esportives; algunes de les quals també passaven de nit. Com ara els partidets de futbol 7 a la Brafa, amb el Marc V, Ruben, David, Jesus, Marc Martinell y Farris he pogut reviure les meves èpoques dorades de futbolista, quasi sempre com a pitxitxi, es clar. També he compartit bons moments fent footing amb el Javi ra ra ra, i el Marc V pel parc Cervantes, malgrat que l'edat de l'un i el genoll de l'altre sovint ho convertien en un passeig.

Es mereixen una menció especial també aquells que amb el seu ajut i suport moral m'han permès superar els últims esculls burocràtics per tal que la tesi arribés a port. Gràcies Farris pels ànims, em sap greu haver-te deixat plantat al juliol; gràcies Ernesto Nicolas per convertir el "vuelva usted mañana" del departament en una pertita en dulce, i sobretot a la Nuria Bayó per haver anat davant meu en tots els tràmits i donar-me consell de primera mà, sense tu no se que hagués fet.

I com que no tenia llit al laboratori, la resta de vida durant aquests anys l'he passada al carrer Camèlies, un gran ¨hostal" amb millors inquilins. Gràcies Esteve per fer-m'hi lloc i guiar-me en el camí de la pseudo-emancipació, gràcies Julito per arrosegar-me al costat fosc de la xarxa i de la societat de la informació, al Kalamar per una visita curta però intensa, i al Paquito-flope-ewok per compartir l'estress final de la tesi, per fotre'm les bronques quan em quedo dormit al sofà, i per deixar-me ser la carabina tants divendres (gràcies Anna). Gràcies també al Miquel alias nailman, que tot i ser una mica ploramiques, ha sigut un gran company de penes, viatges i batallites; confesor, confident, corrector però sobretot amic.

També m'agradaria recordar als amics del poble que han suportat sempre estoicament els llargs períodes sense veure'ns, gràcies: Joan, Esteve, Anna, Genís, Albert, Enric, Lluís ….

A la Judit tot i que sempre em roba les idees, li voldria agrair el temps que hem passat junts -i no tant junts-, els ànims constants, les crítiques benèvoles d'aquest manuscrit, que aguanti tant bé els meus mals humors, les meves paranoies i que em continuï enlluernant amb la seva lucidesa. Ets una de les les millors coses que m'han passat en aquests anys

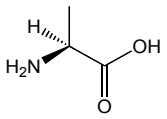
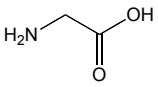
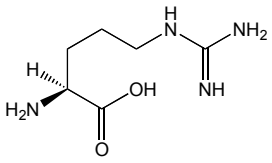
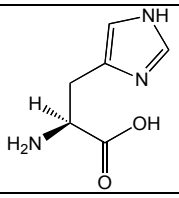
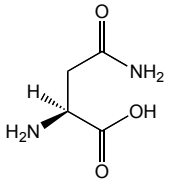
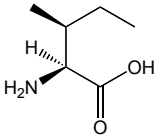
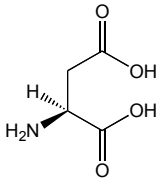
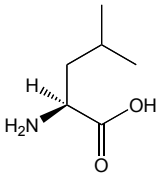
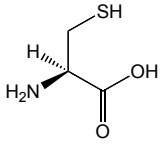
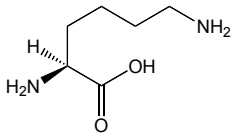
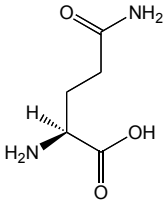
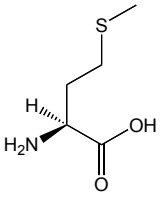
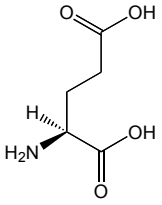
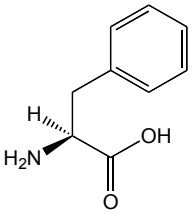
Per acabar m'agradaria donar les gràcies a la meva família pel seu suport incondicional. A la gent de Can Mias per ser-hi sempre que els necessites i per haver-me aguantat i alimentat durant tans anys. Als meus pares, perquè sense ells no seria on sóc, i a la meva germaneta per ser capaç de portar-nos a tots de vòlit.

ABBREVIATIONS

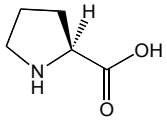
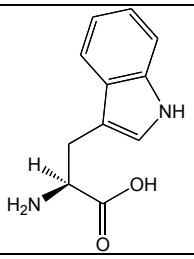
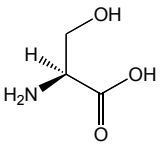
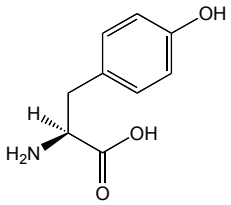
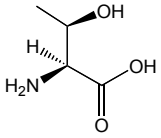
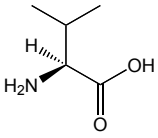
aa	amino acid
AIDS	acquired immunodeficiency syndrome
AIF	apoptosis inducing factor
ATP	adenosine triphosphate
BIR	baculovirus inhibitor of apoptosis protein repeat
COSY	correlation spectroscopy
CPMG	Carr–Purcell–Meiboom–Gill
CS	chemical shift
c.m.c.	critical micellar concentration
CNS	central nervous system
CSA	chemical shift anisotropy
CSP	chemical shift perturbation
DHPC	dihexanoyl phosphatidylcholine
DIABLO	direct IAP binding protein with low pI
DMF	dimethyl formamide
DMPC	dimyristoylphosphatidylcholine
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
DPC	dodecyl phosphatidyl choline
DSA	doxyl stearic acid
DTT	dithiothreitol
EGC	epigallocatechin
EGCG	epigallocatechin gallate
EGF	endothelial growth factor
EGFR	endothelial growth factor receptor
EPL	expressed protein ligation
EPR	electron paramagnetic resonance
ESI-MS	electron spray ionization mass spectrometry
FADD	FAS-associating death domain
FDA	food and drug administration
FGF	fibroblast growth factor
FTIR	Fourier transform infrared spectroscopy
GPCR	G-protein coupled receptor
HIV1	human immunodeficiency virus protein 1
HMBC	heteronuclear multiple bond correlation
HMQC	heteronuclear multiple quantum correlation
HPLC	high performance liquid chromatography
HSQC	heteronuclear single quantum correlation
HTS	high throughput screening
IAA	indoleacrylic acid
IAP	inhibitor of apoptosis protein

IGP	indole-3-glycerol phosphate
ILOE	inter ligand nuclear Overhauser effect
iPrOH	isopropanol
IPTG	isopropyl-beta-D-thiogalactopyranoside
KF	kahalalide F
LDH	lactate dehydrogenase
MALDI-TOF MS	matrix assisted laser desorption ionization time of flight mass spectrometry
MMP	matrix metallo protease
MOA	mode of action
MOM	mitochondrial outer membrane
MRI	magnetic resonance imaging
MS	mass spectrometry
MW	molecular weight
NAD+	nicotinamide adenine dinucleotide
NMR	nuclear magnetic ressonance
NOE	nuclear Overhauser effect
NOESY	nuclear Overhauser effect spectroscopy
PDFG	platelet-derived growth factor
PF	protection factor
POP	prolyl oligo peptidase
PRPP	5-phosphoribosyl-1-pyrophosphate
RDC	residual dipolar coupling
RMSD	root mean square deviation
RNA	ribonucleic acid
ROESY	rotating frame nuclear overhauser effect spectroscopy
SDS	sodium dodecyl sulfate
SEA-TROSY	solvent exposed amides TROSY
SMAC	second mitochondria-derived activator of caspase
STD	saturation transfer difference
TCM	traditional Chinese medicine
TEM	transmission electron microscopy
TFA	trifluoroacetic acid
TFE	trifluoro ethanol
TOCSY	total correlation spectroscopy
TOAC	2,2,6,6-Tetramethylpiperidine-1-oxyl-4-amino-4-carboxylic acid, free radical
TROSY	transverse relaxation optimized spectroscopy
VEGF	vascular endothelial growth factor
VEGFR	vascular endothelial growth factor receptor
WaterLOGSY	water-ligand observed via gradient spectroscopy
XIAP	x-inhibitor of apoptosis protein

ANNEX I: AMINO ACIDS[1]

Name			Structure	Name			Structure
Ala	A	Alanine		Gly	G	Glycine	
Arg	R	Arginine		His	H	Histidine	
Asn	N	Asparagine		Ile	I	Isoleucine	
Asp	D	Aspartic acid		Leu	L	Leucine	
Cys	C	Cysteine		Lys	K	Lysine	
Gln	Q	Glutamine		Met	M	Methionine	
Glu	E	Glutamic acid		Phe	F	Phenylalanine	

[1] Amino-acid abbreviations used follow the rules of the Commission on Biochemical Nomenclature of the IUPAC-IUB as specified in *Eur J Biochem* 1984; **138**: 9-37 and *Eur J Biochem* 1993; **213**: 2.

Pro P Proline 	Trp W Tryptophan 
Ser S Serine 	Tyr Y Tyrosine 
Thr T Threonine 	Val V Valine 

0 INTRODUCTION AND OBJECTIVES

The years where NMR spectroscopy played a mere analytical role in drug discovery and medicinal chemistry programs are long gone. The advent of modern recombinant DNA and protein expression technologies, together with spectrometer hardware evolution have pushed NMR towards the development of an extended list of experiments to characterize large biomacromolecular systems. This arsenal of techniques has seeped into all stages of drug discovery pipelines (Figure 1) triggering a silent revolution; as a result in the following years we should witness a notable boost in the productivity and efficiency of drug development endeavors. [1]

Thus, NMR enjoys healthy adulthood and its collection of experiments supplies a wealth of structural information but also for a wide variety of transient and dynamic processes. Often, it is in this structural sense that NMR contributes to the characterization of therapeutic targets during the very early stages of drug development. Several massive structural genomics projects demonstrate such role in the determination of protein structures; also, and as opposed to X-RAY, NMR provides insight into folding and dynamic phenomena. [2] But even more interesting at this early stages, is the use of NMR experiments to determine protein druggablility; [3, 4] this is to *a priori* assess whether the activity of a given protein could be modulated by organic molecules; such target quality assessment will inevitably increase success rates for the drug development process.

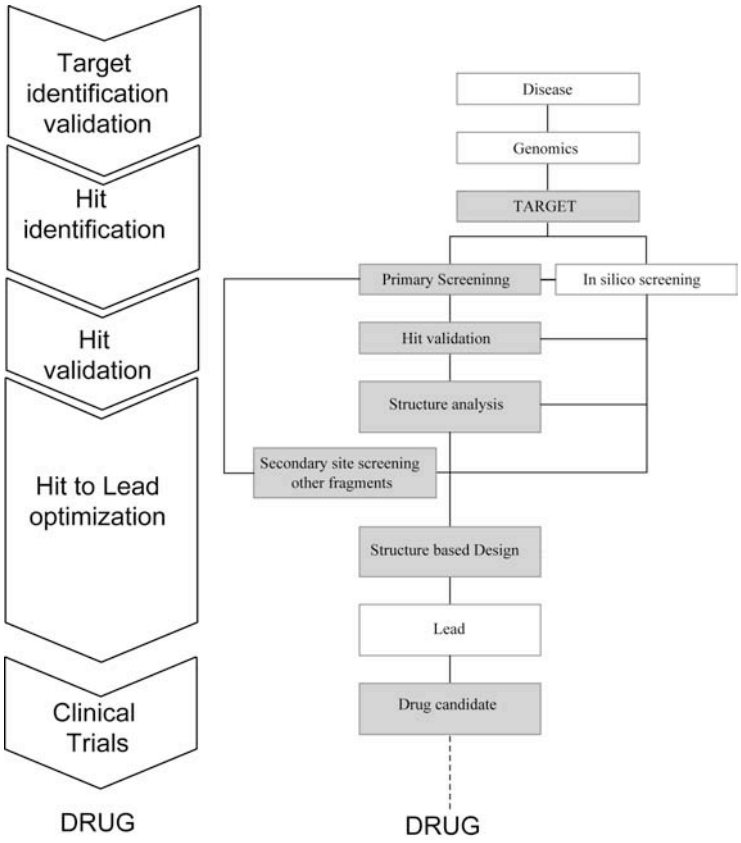


Figure 1 Flowchart outlining the drug discovery process; steps where NMR techniques have found application are shaded.

As we will see in the first part of this introduction, a large portion of these newly developed experiments is devoted to the detection and characterization of molecular recognition events,[5-7] which allow the harvesting of structural, thermodynamic and kinetic information for many intermolecular complexes –either protein-protein or protein-ligand. This wide information range has produced a large impact on drug discovery, providing remarkable improvements to the hit identification process, but also their optimization process, which in the end provides high quality leads and eventually drug candidates.

Other remarkable NMR applications include experiments to assess the overall quality of produced leads and hits. For instance the ALARM NMR approach, developed Hajduk and collaborators, is very useful during hit validation stage to detect false positives associated to reactive molecules.[8] With regard to lead quality, NMR also allows us to anticipate pharmacokinetic and pharmacodynamic properties of leads and drugs particularly by determining their propensity to bind BSA these factors should be taken into account prior to clinical trials.[9]

On a second part of the introduction we will deal with a much more “classic” use of NMR. Structural characterization of bioactive conformations in flexible molecules is another important contribution to the NMR field. Very often medicinal chemists approach the problem of improving flexible bioactive molecules by designing analogs with restricted conformational freedom. When properly designed, these molecules should have a lower entropic penalty towards binding and consequently their affinity and activity should increase. Some relatively new methodologies such as trNOE [10] allow the determination of bioactive conformations for fast exchanging ligands; for stronger binders however, such information is far more difficult to obtain and generally involves determination of the whole complex structure. In the latter case, characterization of the molecular conformational preferences may be very informative, especially when such exercise is performed under different media. Short to medium peptides are usually flexible and they seldom have a stable conformation, thus these molecules serve as valuable test bench to show what NMR has to offer to the conformational analysis of flexible molecules.

0.1 NMR EXPERIMENTS IN DRUG DISCOVERY

There are a variety of parameters sensitive to the intermolecular complex formation event. These include chemical shift, coupling constants, relaxation and translational diffusion properties. In any case, NMR experiments to characterize protein-ligand complexes are typically divided according to which part of the complex is observed: those that detect the interaction by observing the macromolecule and those that measure NMR parameters for the small binding partner.

0.1.1 RECEPTOR BASED EXPERIMENTS

Among the experiments that focus on protein observation to characterize recognition events, probably the simplest (conceptually) is Chemical Shift Perturbation (CSP). This involves monitoring protein resonances upon ligand addition; and is commonly performed with heteronuclear 2D NMR experiments (^{15}N-^{1}H or ^{13}C-^{1}H) on an appropriately labeled protein sample. Chemical Shift is extremely sensitive to environment changes, whether they come from denaturation, pH, temperature or changes associated to ligand binding. Thus perturbations on δHN can be used to assess whether the binding event does occur and its dissociation constant. But the real power of the technique is unleashed when the protein structure and assignment is available. δHN perturbations will occur primarily on resonances nearby the binding site consequently, by plotting the affected set of spins on the structure, one can delineate the protein-ligand interface.

CSP is however not devoid of difficulty; the first limitation is intrinsic to the method and consists on the labeled protein requirement: HSQC protein observation requires stable isotope enrichment most often achieved by means of over expression in E. coli. Nonetheless, these systems do not always succeed when producing eukaryotic proteins that may require posttranslational modifications or even chaperones, and one has to turn to eukaryotic host to obtain active proteins; a solution that is not always practical given the high cost of eukaryotic hosts labeled medium.

But even when E. coli production is possible, cost per sample is still a concern; and although isotope prices have been largely reduced in the last years, the amount of protein required for CSP studies are considerable (~500 μg/sample), especially if one wants to start an HTS program. Fortunately hardware improvements (i.e cryoprobes…) are expected to increase spectrometer sensitivity and consequently reduce protein requirements in the near future.

Another weakness is related to molecular weight range amenable NMR and has a direct incidence on the biological targets that can be tackled with receptor based experiments. Observation of macromolecules or complexes beyond 40 kDa is often hampered by severe relaxation and important signal overlapping. The former effects are characterized by size enhanced signal decays and have important consequences to experiment sensitivities. Recent development of TROSY type experiments combined with protein perdeuteration alleviate the deleterious relaxation effects by compensating CSA and dipole-dipole relaxation mechanisms.[11] This scheme has been a major breakthrough in the field and has expanded the MW range up to 900 kDa in the most favorable conditions. [12]

Another interesting property of TROSY type experiments is they also minimize signal overlapping due to sharper cross-peaks delivered by such schemes; however signal overlapping is still important for very large

systems for which several strategies have been suggested to simplify spectrum complexity while retaining relevant information. SEA-TROSY experiment is one of such examples. [13] It is a modified HSQC-TROSY sequence containing a "SEA" module that selects solvent exposed amides. Protein core amides are filtered out and only the interesting –from an interaction point of view- surface amides are observable in the spectrum. Other strategies achieve simplicity at a protein rather than at a sequence level by carefully placing the desired stable isotopes in a limited number of protein positions. This type of control can be achieved through chemical or semi-synthetic approaches as in Expressed Protein Ligation (EPL) approaches [14] but also by interfering in host's protein biosynthesis through culture media modulation. [15] This second alternative is usually superior in terms of simplicity and yield, and only requires minor changes to protein production protocol thanks to the wide knowledge on bacterial amino acid biosynthesis.

Another major stumbling block for quantitative or semi-quantitative CSP information is protein floppiness, often manifested as large conformational rearrangements occurring in response to a recognition event making it difficult to differentiate between δ perturbations directly caused by the ligand and those induced by distal effects. This limitation is particularly worrisome considering the increasingly important biological role of protein flexibility, which lies at the core of inter-domain cross-talk in proteins, and thus plays and essential role in signaling cascades, protein promiscuity –the ability to interact with multiple ligands– and ultimately determines how protein-protein interaction networks have been constructed in cells. It would seem that the simple key and lock model is obsolete; as a matter of fact some authors suggest flexibility is a general property of protein hot spots. [16, 17]

Several authors have addressed such supple proteins and developed methods to extract structural information. Differential Chemical Shift Perturbation for instance, utilizes closely related small MW ligands, to perform standard CSP titration experiments on the target protein. The assumption here is that similar compounds should bind similarly to protein surface, thus differential changes in the ligand perturbation profiles can be structurally interpreted. This strategy was first used by Fesik [18] and later Wagner et al combined this approach with in silico docking to determine the binding site for a family of small Bcl-XL inhibitors.[19]

Cross-Saturation TROSY is a technique originally developed to determine large protein-protein binding interfaces, for which accurate delineation is difficult due to large size and extensive "induced fit" effects involved in protein complex formation. It was first described by Takahashi and collaborators to study FB-Fc (50 kDa) complex.[20] For this purpose they produced ^{2}H, ^{15}N-labeled FB protein –^{2}H for all non-exchangeable protons-, and Fc partner without any specific labeling, the only restriction being it should contain protons. Sample consisted on FB and Fc (2:1) in 10:90 H_2O/D_2O buffer at pH=6 and the experiment proceeded in two steps: two consecutive ^{15}N-^{1}H TROSY-HSQC type experiments. The first TROSY-HSQC is mainly for control purposes; the second however includes a saturation period on Fc aliphatic resonances. During this period, much as it happens in STD sequences (see below), saturation is spread throughout Fc resonances and complexed FB nearby amides, for which we will observe a signal reduction compared to the control experiment. This saturation transfer phenomenon relies on ^{1}H-^{1}H cross-relaxation, which is solely dependant on inter-proton distance; thus one can interpret a decrease in amide intensity as evidence for involvement in protein-protein interface without fear from conformational artifacts. Of course behind this analysis we assume there is no saturation happening within FB that could account for amide intensity

decrease, for this reason efficient FB perdeuteration is capital to guarantee that Fc partner mediates amide saturation.

0.1.2 LIGAND BASED EXPERIMENTS

Ligand based experiments are more numerous and diverse than receptor based NMR experiments, mainly due to dramatic changes on several small molecule properties when this undergoes binding with a large molecule. Most of these properties are related to the ligand's apparent MW, which in contrast to the free form in the presence of the receptor will be closer to the complex MW; thus size-dependent properties such as longitudinal and transverse relaxation, diffusion coefficients and intra/intermolecular magnetization transfer will report this molecular weight change for the ligand upon complex formation.

Their main advantage over receptor-based experiments is that they don't require the usage of labeled protein; also, they're far less protein demanding and most of the experiments can be performed at 50 μM concentration or below, about one order of magnitude less than receptor-based counterparts. The former is inherent to measuring ligand resonances, the latter however is related to the fact that large ligand/protein ratios are used in these experiments and that the binding effects on the small molecule resonances will be observed throughout the free ligand. Unfortunately binding effects on the free ligand population are only visible when the binding falls in the fast exchange regime, tighter binders like those in slow or intermediate exchange will likely appear as non-binders to the eyes of ligand based experiments, consequently relegating these techniques to weak binder detection. This limitation has been somehow circumvented by the recent development of competition experiments; there, a weak binder is used as a reporter and the presence of tighter binders in the sample is detected by the disappearance of the binding effects on the probe.

Also due to ligand resonances observation, these techniques have no upper molecular weight limit on the proteins they can deal with, in fact the higher the molecular weight the more important will the effects on the ligand be, and the easier to the detect; in these sense Meyer and col. [21, 22] reported solid support immobilization of receptors as a method to boost their experiments performance and allowing the authors to reutilize the receptor, and the same authors even provide examples where the assay has been performed on membrane proteins using whole cells.[23] It would seem that ligand based experiments have been released from most protein limitations imposed to Chemical Shift Perturbation; but this is not free and the price paid is structural information, since little or no information regarding the binding mode can be obtained. For these reasons, and as we'll see later most of these experiments are often found at early stages of drug discovery programs: i.e. high throughput primary screenings.

0.1.2.1 RELAXATION MEASUREMENTS

Relaxation properties for the free and bound forms of a ligand are quite different, as a consequence, the measured R_1 and R_2 values for a ligand undergoing binding are a weighted average accounting for the free and bound relative populations; thus the relaxation enhancement suffered by a binding molecule upon complexation can be used to detect such event.

On the one hand T_1 relaxation rates to detect binding can only be used when such measurements are performed on individual resonances, this is due to the relaxation mechanisms involved with multiple resonances; this has disastrous implications when dealing with compound mixtures and consequently limits the scope of the experiment for screening purposes.

Conversely the use of T_2 relaxation for screening purposes has been much more successful. Relaxation times for small molecules or rapidly tumbling are usually long, while for resonances in large molecules or protein-bound ligands are usually much shorter. This difference is usually translated into apparent differences in their resonance linewidth, which has a direct dependence on the transversal relaxation rate (R_2/π). Thus, in the presence of the protein a ligand will see its resonances immediately broadened; and such effect can be used to pin down binding within a mixture of molecules. Alternatively, this effect can be magnified by using T_2-filters (CPMG) that will allow ready identification of compounds with shorter relaxation times. [24] Similarly, Jahnke and collaborators have reported increased experiment sensitivity by covalently attaching a paramagnetic probe to the protein surface; the presence of unpaired electrons on the surface further enhances transversal relaxation for binding compounds and makes it easier to distinguish between ligands and non-ligands.[25]

0.1.2.2 ^{19}F Experiments

^{19}F nucleus has lately attracted much attention in the screening field for various reasons. First and most important is its high sensitivity; this is due to ^{19}F's high gyromagnetic constant ($\gamma_F \sim 0.94\ \gamma_H$) but also to its 100% natural abundance. Moreover, fluorine displays a wide chemical shift range, which makes it unlikely for overlapping to occur. Other useful properties include its high sensitivity to binding events; this often induces important changes to fluorine chemical shift and linewidth due to fluorine's large CSA. Given the remarkable relaxation properties above, binding detection can often be performed simply by acquiring 1D spectra in absence and in the presence of protein target. Nonetheless, a much more general approach to discriminate between fluorine containing ligands among non-binders consists on the determination on transversal relaxation rates (T_2) –either by direct measurement of linewidth or by using T_2 filters- much in the same way as in 0.1.2.1. [26, 27]

Obviously such experiments are limited to some extent, as they require the presence of fluorine atoms in the assayed compounds. This atom however, is rather common in commercial drugs and often medicinal chemists enhance metabolic stability or modulate compound lipophilicity by incorporating such atoms. On the other hand, fluorine is seldom found in natural sources or solvents; for these reasons it has been proposed as an alternative to avoid background resonances arising from the protein or the annoying solvent signals, which often have deleterious consequences for the performance of ligand based experiments.

0.1.2.3 Magnetization Transfer

Probably the most successful and remarkable experiments in this class are STD [22, 28] and WaterLOGSY [29]. Both take advantage of the strong negative intermolecular nOe that is characteristic of large receptor-ligand complexes. In both cases non-equilibrium magnetization or saturation is created on the complex and, thanks to the strong negative nOe, transferred to the ligand that ultimately suffers perturbations on the resonance intensities of its free form.

In the case of STD experiment, non-equilibrium magnetization is achieved by selective excitation (on-resonance) of target resonances; spin diffusion eventually transfers magnetization to the bound ligand. When the ligand dissociates from the target into solution, the magnetization change transferred in the bound state is retained in the free ligand. The difference spectrum between the on-resonance experiment and a

reference spectrum, acquired with off-resonance irradiation, yields a spectrum containing only those ligand signals that have been perturbed by binding to the target. (Figure 2)

As for WaterLOGSY, the experiment is very similar, only that irradiation is performed on water instead. Eventually magnetization is transferred from this large reservoir to the protein surface through different mechanisms, including chemical exchange and hydration water. Later spin diffusion relays magnetization to the ligand providing similar effects as STD for binding molecules. Non-binders however, will suffer a negative cross-saturation effect that in some instances can induce to errors.

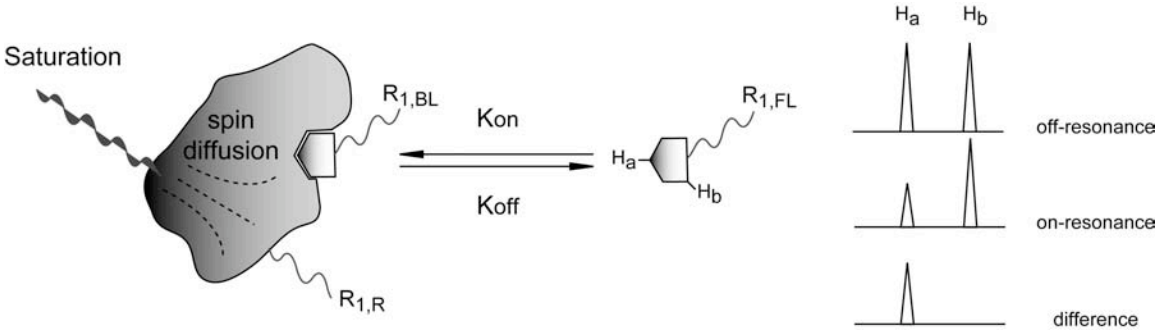


Figure 2 Illustration for STD experiment. The resulting non-equilibrium magnetization from on-resonance irradiation spreads out across the target molecule and is transferred to the ligand, especially to close-by protons. The resulting on and off-resonance experiments are depicted, and the difference experiment both asserts binding and allows us to identify ligand protons in closer contact with the protein. The magnetization of the bound ligand decays rapidly with a rate of $R_{1,BL}$, which is of the same order as the rate for the target molecule $R_{1,R}$. On dissociation into solution, the relaxation properties of the ligand change and the acquired non-equilibrium magnetization now decays with a rate of $R_{1,FL}$ for the free ligand.

0.1.2.4 COMPETITION EXPERIMENTS

As already mentioned, one of the major drawbacks associated with ligand based experiments is their limited affinity detection range, which usually confines detectable binders within the μM range The reason being long residence times for strong binders in the binding pocket; this translates into slow exchange in the NMR timescale and as a consequence information on the bound state is easily lost due to the adverse relaxation properties for the complex.

In order to circumvent such limitations several competition-based screening approaches have been proposed.[30-32] In general, such strategies require knowledge of a weak binder that can be detected with the given ligand based experiment; this compound functions as a reporter. The presence of a tighter molecule in the sample, targeting the same binding site, will prevent observation of binding effect on the reporter molecule. Such indirect detection relies solely on the reporter resonances and involves the need for deconvolution when working with compound mixtures. Nonetheless, this limitation is clearly outweighed when we consider the wealth of information provided by competition experiments, since K_D and binding site characterization for the reporter molecule will provide a means to estimate both affinity and binding mode for the competing molecule.

In principle, all ligand-based techniques can be extended to a competition type of experiment. However, experiments using simple 1D spectra are especially worth considering: WaterLOGSY, STD, and ^{19}F-screening. Probably the most attractive is the competition-based fluorine screening, for it combines the high sensitivity of its relaxation parameters with ^{19}F clean spectra devoid of solvent and target signals.

0.1.3 ENHANCED DRUG DISCOVERY EMPLOYING NMR

The various technologies presented in the previous section allow NMR to participate nearly on every stage of the drug discovery process: Ligand based experiments contribute to lead identification thanks to their throughput and low cost, and although during early design stages receptor based techniques only provide limited structural information, complete structure is available for the macromolecular complexes with tight ligands, as it is routinely done for proteins. Despite each particular experiment's limitations it is undeniable that NMR as a technique is very useful in the drug discovery field. It is not surprising to see almost every pharmaceutical company trying to muster its power into their drug design pipelines; some even developing completely NMR-guided drug design strategies. In this sense a very attractive combination consists on the use of NMR and fragment-based lead discovery strategies.

As opposed to traditional HTS, which rely on massively assaying complex chemical libraries in order to identify potent leads, a more efficient approach has been lately gaining interest. Instead, screening is performed on libraries of so called fragments or shapes; these are low MW molecules with few chemical functions and thus expectedly less affinity towards the protein target (mM-μM) than a typical lead. Positive compounds resulting from the screening can be regarded as simple binding epitopes and will be later evolved into more conventional lead molecules through various stages.

Bottom-up lead building offers several advantages over traditional HTS schemes. First the number of molecules to be tested is small, in the order of hundreds compared to the thousands for massive screening assays, due to low complexity for the tested fragments. Also in terms of efficiency, fragment screening is superior to other massive strategies; although the resulting positive fragments may present weak affinities towards the protein target, their binding is usually very efficient when affinity per mass unit is considered and compared to HTS hits these fragments have a large fraction of their atoms interacting with the receptor.

Fragment screening can be performed using various methods, however X-RAY and NMR [6, 33] are especially suitable to detect interactions in the μM to mM range; but more importantly once positive fragments have been identified both techniques provide binding structural validation.

On the grounds of such X-RAY or NMR information, hypotheses can be performed as to which chemical modifications on the initial hit will provide affinity increases; examples of this kind of fragment evolution are rather common in the literature leading to very potent molecules after several cycles of structure guided evolution.

Fragment linking is another attractive way to evolve positive shapes provided by screening. This approach can only be performed when two fragments have been identified to be sitting on adjacent protein binding sites, and it consists on designing a favorable chemical linkage between the two fragments to produce a bi-ligand with increased potency. Its enhanced affinity does not only stem from additive fragment interactions but more importantly from an entropic component since the freedom-loss penalty associated with bi-ligand binding will be less than the sum of the penalties for all individual fragments. (see Equation 1)

$$\Delta G^{AB} = \Delta H^{A} + \Delta H^{B} - \Delta T \cdot S^{AB} = R \cdot T \ln K_D{}^{AB}$$

Equation 1

This approach is also very design intensive and relies heavily on the structural information obtained for interacting fragments, thus in such strategy X-RAY or NMR play a capital role in guiding the linker design. For the latter, ILOE or CSP experiments are in this regard very helpful. [6, 34, 35]

Once a fragment-based strategy has been chosen over a traditional HTS approach, another important issue has to be addressed: building a fragment chemical library. HTS libraries usually follow a series of guidelines in their compound composition so as to maximize their "drug-likeness" and in this way increase the probability for the resulting leads to survive the development process and become orally bioavailable drugs. Probably the most typical guidelines used to build chemical libraries are the "Lipinski rule of five" [36], which basically establish a range of recommended values for various properties (molecular weight, octanol/water partition coefficient, hydrogen bond acceptors/donors, number of rotatable bonds…) for a molecule to be suitable for oral administration. Observation of several successful and failed stories has produced a similar set of rules to define fragment-like compounds, and parallel to the Lipinski recommendations these have been called the "rule of three". In general, such guidelines are based on the fact that a typical lead optimization procedure significantly increases lipophilicity and molecular weight; consequently Lipinski thresholds for lead-like and fragment-like molecules should be reduced to account for fragment-to-lead and lead-to-drug development.

According to this, the attributes for a fragment-like molecule would be: MW<300 (usually above 100 uma), a logP below 3, less than 3 hydrogen bond acceptors and 3 hydrogen bond donors, and less than 3 rotatable bonds. All in all, the motto behind quality fragment libraries is "simple is beautiful", simple fragments will produce quality leads with room for improvement in their lipophilicity, MW, etc.

0.2 PEPTIDE NMR

Compared to the relatively compact nature of most proteins, small peptides still represent a major challenge in structural characterization due to coexistence of many conformations. With varying degrees of success, often depending on the size of the molecule, well-folded proteins can have their structure determined either by X-Ray or NMR. Here, densely packed folds keep atoms in relatively fixed positions, which in turn yield reliable experimental constraints to guide structure calculations.

However the scenario changes when one studies partially folded proteins or small peptides. In the former, X-Ray is outperformed, due to difficult crystallizations or artifacts in the resulting structures; but also because NMR relaxation measurements offer dynamic information to build more "realistic" models. The latter though is more complicated given the cost of labeled peptide production, especially when these molecules include non-proteinogenic modifications. In such instances very limited structural information builds fuzzy models able only to describe conformational preferences.

Structural information for small peptides has traditionally been obtained from circular dicroism, Fluorescence, FTIR, Raman, and also NMR; which in general are more or less complementary. Although all techniques are able to time-resolve conformational exchange phenomena, they all have limitations. Fluorescence requires the introduction of probes to the system, Raman and FTIR spectroscopy lack enough frequency dispersion for residue specific information; and circular dicroism is also positionally averaged due to its dependence on cooperative amide transitions. As a result, NMR is probably the most powerful technique in the list, due to its large number of parameters sensitive to conformational changes, and also its frequency dispersion that allows us to get the information at residue or atomic resolution.

But prior to the structural information, one must go through the resonance assignment process, which for non-labeled peptides proceeds through acquisition of complementary 2D homo-nuclear experiments, as described by Wütrich and co-workers [37]. The first experiments (2D-TOCSY or 2D-COSY) allow the identification of resonances connected through ^{1}H-^{1}H J couplings; particularly for peptides this means within a residue. Once these spin systems –or sets of J connected protons- have been identified and even associated to a particular amino acid type, 2D-NOESY or 2D-ROESY complementary experiment is used to determine inter-residue connectivity given their ability to correlate spatially close protons.

But beyond its use in peptide assignment, nuclear Overhauser effect and NOESY experiments are probably the most widely used NMR parameter for biomolecule structural, and conformational analysis. And before the recent development of Residual Dipolar Coupling techniques, protein and peptide structure calculation was limited to the nOe effect as the only source of long-range structural information.

0.2.1.1 nOE

The nOe is a phenomenon that arises from the cross-relaxation between two nuclei; there are a number of mono and multi-dimensional NMR experiments to observe this effect. In any case, the intensity of the nOe is a function of the distance (proportional to r^{-6}) between the cross-relaxing nuclei. In proteins though, highly crowded proton environments often trigger non-linear nOe behaviors due to the presence of other cross-relaxation pathways and care must be taken to avoid artifacts in the structure calculation.

Conversely, in peptides, the major problem is flexibility. Intermolecular rearrangements average out long-range effects and leads to the prevalence of short and medium-range nOes, which may merely be representative of a small population of peptide conformations. Thus, conformational averaging tends to bias nOe values, which frequently translates in over-emphasized structural tendencies.

Yet other limitations to the use of nOes, especially with small peptides, are its sign and intensity. Since as a result of their size and correlation time peptide resonances tend to display weak negative nOes; or even under some instances different protons within the same peptide may experience positive and negative nOes respectively, depending on the conformational averaging taking place and the resulting correlation time of each inter-proton vector responsible for the nOe.

0.2.1.2 SPIN-SPIN COUPLINGS

Another parameter traditionally used in conformational analysis by NMR are vicinal coupling constants [38] whose dependence on dihedral angle values is expressed by the Karplus relationship.(Equation 2)

Particularly for peptides and proteins there are various useful coupling constants with potential information on ϕ, $\psi$and χ dihedral angles: $^{3}J^{13}C'\text{-}NC_{\alpha}\text{-}^{1}H$, $^{3}J^{1}H\text{-}NC_{\alpha}\text{-}^{13}C_{\beta,}$etc. For non-labeled peptides however, the easiest measurable constants are $^{3}J\alpha N$ and $^{3}J\alpha\beta$, either through 1D ^{1}H or 2D ^{1}H homo-correlation E-COSY. Once measured, they can be correlated to ϕ and χ_1 angles respectively in the peptide; yet due to Karplus curve intrinsic ambiguity it is difficult to extract a unique result, without the evaluation of additional coupling constants.

$$^{3}J_{H\alpha N} = 6.7\cos^{2}(\phi - 60) - 1.3\cos(\phi - 60) + 1.5$$
$$^{3}J_{\alpha\beta} = 9.4\cos^{2}\chi_1 - 1.4\cos\chi_1 + 1.6$$

Equation 2: Karplus expressions correlating ϕ an χ_1 angles in peptides to $^{3}JH\alpha N$ and $^{3}J\alpha\beta$ coupling constants respectively.

Added to the intrinsic ambiguity in Karplus relationship, conformational averaging also has deleterious effects as those described for the nOe effect. For both dihedral angles ($\phi$and χ) energy minima correspond to extreme values of $^{3}J\alpha N$ and $^{3}J\alpha\beta$ coupling constants. This means that when an exchange event occurs between conformations situated in these minima the measured coupling constant will be an average value, void of significant structural information.

0.2.1.3 CHEMICAL SHIFT DIFFERENCE

Chemical shifts are the most readily accessible quantities in NMR and, because its particular value for a given nuclei depends on a variety of structural factors, they have traditionally been regarded as very rich source of information.

Several practical issues have traditionally threatened however the quality of this information, in the beginning this was related to the relative sparseness of NMR data, and later to the lack of standardization in calibrating methods throughout the scientific community. But beyond these limitations chemical shift information has been used to evaluate and characterize the secondary structure content in peptides, analyze conformationally flexible peptides, and even refine protein tertiary structures.

The large body of assignments and theoretical works around these have shed some light to the effects responsible for chemical dispersion in peptides and proteins. For peptides it is recognized that the magnetic anisotropy of the peptide group is an important contributor to chemical shifts in proteins, which explains to a large extent the sensitivity of Hα protons to local secondary structure. Other effects include the ring current present in the vicinity of aromatic amino acids or shifts induced by electrostatic or dipole effects.

Typically the procedure behind chemical shift analysis is as follows. We first need a reference chemical shift; in this regard, what is known as Random coil chemical shifts, or those expected for a non-structured peptides have been proposed by various authors [38]. These values for the different nuclei in each amino acid (Hα, ^{13}Cα, Hβ, H_N, ^{15}N, $^{13}C_O$...) have in general been determined using short oligo-peptides. The analysis proceeds by comparing the actual chemical shift for a given resonance in the peptide or protein to its corresponding tabulated random coil value. The differences are often analyzed in terms of secondary structure participation, for instance if a particular amino-acid displays a Hα biased towards higher fields it is attributed to be involved in α-helical structures, on the contrary if Hα shifts fall to lower fields the residue would participate in a β-strand.

For the particular case of non-labeled amino acids, Hα is usually the most useful resonance to study, several authors however suggest the usage of δNH to characterize helical secondary structures, especially to detect hydrogen bond pattern distortions in amphipathic helices [38, 39]. Conformational exchange is also an issue, and as for spin-spin couplings extreme chemical shifts observed in α-helices or β-strands are often averaged into values close to the random coil, which again are difficult to interpret structurally.

In our case, the presence of D-amino acids, along with other non-proteinogenic modifications in the peptide will hamper the structural interpretation of chemical shifts. Since, as mentioned before, Hα chemical shift is predominantly dominated by peptide group anisotropy, which will expectedly vary between D and L amino acids, and for this reason tabulated random coil values will likely not be applicable in our case.

0.2.1.4 HYDROGEN BONDS

Another parameter routinely used for structure calculation is hydrogen bond formation by protein or peptide amides. Slow water exchange rates for these protons are evidence for hydrogen bond formation, and can be introduced as restriction in late stages of structure calculation.

In peptides however, direct exchange rate measurements are seldom performed. Exchange rates are often too fast to be measured in water or methanol, conversely in organic solvents such as DMSO exchange does not occur. Instead measurement of NH chemical shift dependence with temperature has been used as an indication of hydrogen bond formation. This is especially true for hydrogen bonding solvents such as DMSO, H_2O and methanol.[40] In DMSO, temperature induced amide shifts below the 4 ppb/K threshold have been considered to be hydrogen bond indication.

These effects can be interpreted in terms of amide exposure to the solvent, since intermolecular hydrogen bonds are readily cleaved with temperature, especially those involving solvent molecules. According to several authors hydrogen bond involvement can also be diagnosed to some extent by the amide chemical shift, although this alternative is not free from artifacts.

All these parameters are also sensitive to conformational exchange, and as in the previous cases caution is advised especially since exchange phenomena involving conformations with different amide solvent exposure could lead to underestimation of hydrogen bonding amides.

0.3 OBJECTIVES

Given the increasing interest raised by NMR in the drug discovery field as well as the structural challenge still posed by medium and small flexible peptides the objectives for the present PhD thesis will be the following:

- In a first stage we intend to **get acquainted with several of its most outstanding methodologies** for the detection and characterization of binding events, including both **ligand** and **receptor based experiments** and using several cancer related protein targets. It will also be interesting to see how these methods are integrated in **holistic drug discovery approaches**.
- In a second stage the previous NMR technologies will be applied to the **development of protein-protein interaction inhibitors** in particular of the pro-angiogenic protein **VEGF**. Several milestones will have to be covered prior to the application of the various NMR screening and design methods, mainly protein production, ligand design and synthesis.
- **Structural characterization of medium sized flexible oligopeptides by NMR**. Exploration of their conformational propensities in different media and structural modelizing using NMR restraints. In particular we will focus our efforts in the **characterization of Kahalalide F** a marine origin depsipeptide with very promising anticancer activity.

0.4 BIBLIOGRAPHY

1. Widmer, H. and W. Jahnke, *Protein NMR in biomedical research.* Cell Mol Life Sci, 2004. **61**(5): p. 580-99.
2. Dyson, H.J. and P.E. Wright, *Elucidation of the protein folding landscape by NMR*, in *Nuclear Magnetic Resonance Of Biological Macromolecules, Part C*. 2005. p. 299-+.
3. Hajduk, P.J., J.R. Huth, and C. Tse, *Predicting protein druggability.* Drug Discov Today, 2005. **10**(23-24): p. 1675-82.
4. Hajduk, P.J., J.R. Huth, and S.W. Fesik, *Druggability indices for protein targets derived from NMR-based screening data.* J Med Chem, 2005. **48**(7): p. 2518-25.
5. Hajduk, P.J., R.P. Meadows, and S.W. Fesik, *NMR-based screening in drug discovery.* Q Rev Biophys, 1999. **32**(3): p. 211-40.
6. Pellecchia, M., D.S. Sem, and K. Wuthrich, *NMR in drug discovery.* Nat Rev Drug Discov, 2002. **1**(3): p. 211-9.
7. Salvatella, X. and E. Giralt, *NMR-based methods and strategies for drug discovery.* Chem Soc Rev, 2003. **32**(6): p. 365-72.
8. Huth, J.R., et al., *ALARM NMR: a rapid and robust experimental method to detect reactive false positives in biochemical screens.* J Am Chem Soc, 2005. **127**(1): p. 217-24.
9. Sun, C.H., J.R. Huth, and P.J. Hajduk, *NMR in pharmacokinetic and pharmacodynamic profiling.* Chembiochem, 2005. **6**(9): p. 1592-+.
10. London, R.E., *Theoretical analysis of the inter-ligand overhauser effect: a new approach for mapping structural relationships of macromolecular ligands.* J Magn Reson, 1999. **141**(2): p. 301-11.
11. Pervushin, K., et al., *Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution.* Proc Natl Acad Sci U S A, 1997. **94**(23): p. 12366-71.
12. Fiaux, J., et al., *NMR analysis of a 900K GroEL-GroES complex.* Nature, 2002. **418**(6894): p. 207-211.
13. Pellecchia, M., et al., *SEA-TROSY (solvent exposed amides with TROSY): a method to resolve the problem of spectral overlap in very large proteins.* J Am Chem Soc, 2001. **123**(19): p. 4633-4.
14. Dawson, P.E. and S.B. Kent, *Synthesis of native proteins by chemical ligation.* Annu Rev Biochem, 2000. **69**: p. 923-60.
15. Weigelt, J., et al., *Site-selective labeling strategies for screening by NMR.* Comb Chem High Throughput Screen, 2002. **5**(8): p. 623-30.
16. Luque, I., S.A. Leavitt, and E. Freire, *The linkage between protein folding and functional cooperativity: two sides of the same coin?* Annu Rev Biophys Biomol Struct, 2002. **31**: p. 235-56.
17. Luque, I. and E. Freire, *Structural stability of binding sites: consequences for binding affinity and allosteric effects.* Proteins, 2000. **Suppl 4**: p. 63-71.
18. Medek, A., et al., *The use of differential chemical shifts for determining the binding site location and orientation of protein-bound ligands.* Journal Of The American Chemical Society, 2000. **122**(6): p. 1241-1242.
19. Lugovskoy, A.A., et al., *A novel approach for characterizing protein ligand complexes: molecular basis for specificity of small-molecule Bcl-2 inhibitors.* J Am Chem Soc, 2002. **124**(7): p. 1234-40.
20. Takahashi, H., et al., *A novel NMR method for determining the interfaces of large protein-protein complexes.* Nat Struct Biol, 2000. **7**(3): p. 220-3.
21. Mayer, M. and B. Meyer, *Group epitope mapping by saturation transfer difference NMR to identify segments of a ligand in direct contact with a protein receptor.* J Am Chem Soc, 2001. **123**(25): p. 6108-17.
22. Mayer, M. and B. Meyer, *Characterization of ligand binding by saturation transfer difference NMR spectroscopy.* Angewandte Chemie-International Edition, 1999. **38**(12): p. 1784-1788.

23. Claasen, B., et al., *Direct observation of ligand binding to membrane proteins in living cells by a saturation transfer double difference (STDD) NMR spectroscopy method shows a significantly higher affinity of integrin alpha(IIb)beta3 in native platelets than in liposomes.* J Am Chem Soc, 2005. **127**(3): p. 916-9.
24. Hajduk, P.J., E.T. Olejniczak, and S.W. Fesik, *One-dimensional relaxation- and diffusion-edited NMR methods for screening compounds that bind to macromolecules.* Journal Of The American Chemical Society, 1997. **119**(50): p. 12257-12261.
25. Jahnke, W., S. Rudisser, and M. Zurini, *Spin label enhanced NMR screening.* J Am Chem Soc, 2001. **123**(13): p. 3149-50.
26. Dalvit, C., et al., *Sensitivity improvement in 19F NMR-based screening experiments: theoretical considerations and experimental applications.* J Am Chem Soc, 2005. **127**(38): p. 13380-5.
27. Dalvit, C., et al., *Fluorine-NMR experiments for high-throughput screening: theoretical aspects, practical considerations, and range of applicability.* J Am Chem Soc, 2003. **125**(25): p. 7696-703.
28. Klein, J., et al., *Detecting binding affinity to immobilized receptor proteins in compound libraries by HR-MAS STD NMR.* Journal Of The American Chemical Society, 1999. **121**(22): p. 5336-5337.
29. Dalvit, C., et al., *WaterLOGSY as a method for primary NMR screening: practical aspects and range of applicability.* J Biomol NMR, 2001. **21**(4): p. 349-59.
30. Jahnke, W., et al., *NMR reporter screening for the detection of high-affinity ligands.* Angew Chem Int Ed Engl, 2002. **41**(18): p. 3420-3.
31. Dalvit, C., et al., *NMR-Based screening with competition water-ligand observed via gradient spectroscopy experiments: detection of high-affinity ligands.* J Med Chem, 2002. **45**(12): p. 2610-4.
32. Dalvit, C., et al., *High-throughput NMR-based screening with competition binding experiments.* J Am Chem Soc, 2002. **124**(26): p. 7702-9.
33. Lepre, C.A., J.M. Moore, and J.W. Peng, *Theory and applications of NMR-based screening in pharmaceutical research.* Chem Rev, 2004. **104**(8): p. 3641-76.
34. Becattini, B. and M. Pellecchia, *SAR by ILOEs: An NMR-Based Approach to Reverse Chemical Genetics.* Chemistry, 2005.
35. Shuker, S.B., et al., *Discovering high-affinity ligands for proteins: SAR by NMR.* Science, 1996. **274**(5292): p. 1531-4.
36. Lipinski, C.A., et al., *Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings.* Adv Drug Deliv Rev, 2001. **46**(1-3): p. 3-26.
37. Wagner, G., A. Kumar, and K. Wuthrich, *Systematic application of two-dimensional 1H nuclear-magnetic-resonance techniques for studies of proteins. 2. Combined use of correlated spectroscopy and nuclear Overhauser spectroscopy for sequential assignments of backbone resonances and elucidation of polypeptide secondary structures.* Eur J Biochem, 1981. **114**(2): p. 375-84.
38. Case, D.A., H.J. Dyson, and P.E. Wright, *Use of chemical shifts and coupling constants in nuclear magnetic resonance structural studies on peptides and proteins.* Methods Enzymol, 1994. **239**: p. 392-416.
39. Orner, B.P., et al., *De novo protein surface design: use of cation-pi interactions to enhance binding between an alpha-helical peptide and a cationic molecule in 50 % aqueous solution.* Angew Chem Int Ed Engl, 2002. **41**(1): p. 117-9.
40. Horst, K., *Conformation and Biological Activity of Cyclic Peptides.* Angewandte Chemie International Edition in English, 1982. **21**(7): p. 512-523.