UNIVERSITA' DEGLI STUDI
DI MILANO – BICOCCA

SCUOLA DI DOTTORATO DI SCIENZE

Facoltà di Scienze Matematiche, Fisiche e Naturali

Corso di Dottorato di Ricerca in Biotecnologie Industriali XXV ciclo

Characterization of Biomedical Relevant Ligand-Protein Interactions Using Nuclear Magnetic Resonance (NMR) Spectroscopy

Tutor: Prof. Francesco Nicotra
Co-Tutor: Dr. ssa Cristina Airoldi

Matr.059161
Silvia Merlo

Dipartimento di Biotecnologie e Bioscienze

Index

Introduction and General Background

1. Nuclear Magnetic Resonance: A Brief And Simplified Overview

Since its discovery in 1945 by Purcell, Torrey, and Pound at Harvard University and by Bloch, Hansen, and Packard at Stanford University, Nuclear Magnetic Resonance (NMR) spectroscopy is one of the most powerful techniques for studying the structure, function and dynamics of biological macromolecules.[1]

Similar to all the other forms of spectroscopy, NMR spectra arise from transitions made by atomic nuclei between different energy states (indeed, this is an oversimplification; for more details, see for example Keeler).[2] Nevertheless NMR differs from the other spectroscopic techniques in three important ways: (1) it looks at how the nuclei of a specific, user-selected, chemical element are distributed amongst the molecules of a sample; (2) NMR signals are sensitive to the local surroundings of the nuclei under observation, providing a tool that can probe the chemical and physical environment of an atom and which can reveal more information about a given sample than the other spectroscopic techniques; (3) it is more highly penetrating—but, usefully, less damaging—than other forms of spectroscopy.[3]

NMR is based on the phenomenon that the nuclei of many isotopes, such as ^{1}H, ^{13}C, ^{15}N and ^{31}P, carry magnetic dipoles. These dipoles take up different orientations in a magnetic field, such as the magnet of an NMR spectrometer, and each orientation has a different energy. Transitions between states with certain energies are permitted according to the postulates of quantum mechanics and, when pulses of electromagnetic

[1] Kwan A.H.*et al.* (2011) FEBS Journal *278*, 687–703.
[2] Keeler J. (2005) published by Wiley & Sons.
[3] Bothwell J.H.F. and Griffin J.L. (2011) *Biol. Rev. 86*, 493–510.

radiation at frequencies that precisely match these energy gaps are applied, it is possible to observe transitions that give rise to NMR signals. As already mentioned, nuclei in different chemical environments (for example the different ^{1}H nuclei in a protein) will resonate at different frequencies and a plot of intensity against resonance frequency is known as a 1D NMR spectrum. Resonance frequencies are typically reported as 'chemical shifts' in units of p.p.m., which corrects for the fact that the raw frequencies (usually in units of MHz) scale with the size of the NMR magnet. One of the key features of liquid state NMR is that the excited states are relatively long lived, with lifetimes in the millisecond–second range (in contrast to the nanosecond timescales that define fluorescence or infrared spectroscopy). Consequently, it is possible to manipulate the excited state to pass excitation from one nucleus to another and, indeed, multiple transfer steps are common in a single experiment. Because NMR allows to measure the frequencies of each of the nuclei through which excitation (magnetization) is passed, signals that correlate (link) the frequencies of two, three or more nuclei can be obtained. In such correlation spectra, each transfer can be visualized as an independent nuclear frequency dimension (axis) and signals occurring at the intersection of two or more frequencies indicate a correlation between the corresponding nuclei. The resulting multidimensional spectra allow to determine unambiguously which signal in a spectrum arises from which atom in the molecule. This process of frequency assignment is an essential step in extracting structural and functional information about the system under study.[1]

2. NMR As A Tool To Probe Molecular Recognition Events

It is a general assumption that the biological function of a macromolecule rely on its interactions with ligand molecules and, as a consequence, the precise characterization of the molecular recognition events is crucial to have a deeper understanding of the mechanisms underlying phenomena of biological and biomedical relevance.[4] In particular, for a macromolecule such as a protein or a nucleic acid, a ligand can be represented by another macromolecule, for example activator and effector proteins, DNA or RNA, lipids, or by a small molecule, for example a second messenger, a cofactor, an inhibitor, a drug.

There have been many attempts in order to accurately characterize macromolecule binding processes with a variety of biophysical methods. Among these methods, NMR has found wide application, as it is a reliable and versatile tool for the investigation of structural and dynamic aspects of biological molecules and their interactions in solution providing structural information at atomic level. This knowledge can also be exploited for the rational design of new therapeutic and diagnostic strategies.[4]

Moreover NMR has the advantage to examine the sample in solution, in a condition quite similar to the physiological one, and the results obtained take into account the flexibility of the protein and the conformational changes that can occur upon ligand binding.

More recently it has also become possible to perform studies also on samples that are partially or completely insoluble in physiological conditions, such as systems containing living cells, nanoparticles, biocompatible materials employed for tissue regeneration etc. This

[4] Meyer B. and Peters T.(2003), Angew. Chem. Int. Ed. *42*, 864-890.

prevents the use of high resolution NMR techniques in solution, and requires HR-MAS (High Resolution-Magic Angle Spinning) probes. This technique allows to obtain high resolution spectra also on partially dissolved samples, making them spinning at the magic angle, and therefore can be exploited for the study of recognition processes in heterogeneous conditions (see chapter **4**).

2.1 NMR Techniques For Ligand Screening

Aim of this chapter is not to make a complete discussion on the most important existing techniques employable for protein interaction studies, but to provide information for the understanding of the applications described in the present thesis.

When the ligand is a small molecule, a lot of different experiments can be performed, by monitoring several ligand NMR variables. They can be classified depending on whether the ligand resonances or the target resonances are observed; the most convenient approach is chosen essentially depending on the possibility to express and purify the receptor, its dimensions, and the interaction strength characterizing the complex. Both the approaches present characteristic features and advantages and disadvantages hereinafter briefly described. The most important techniques employed for both the ligand and the receptor-based approaches are summarized in Table **1** and qualitatively described in the following detailed discussion.

Table 1. NMR spectroscopy techniques for the identification and characterization of ligand binding to a receptor of interest.

Method	Limits and requirements			Identification of		
	Target MW limit	Affinity limit	Target labelling required	Binding site on target	Binding epitope on ligand	Ligand(s) in mixture
STD	lower	upper lower	no	no	yes	yes
Water LOGSY	lower	upper lower	no	no	yes	yes
trNOE	lower	upper lower	no	no	no	yes
Reverse NOE pumping	lower	upper lower	no	no	yes	yes
Relaxation filtering	lower	upper lower	no	no	no	yes
Diffusion filtering	lower	upper lower	no	no	no	yes
SAR by NMR $^{13}C^{15}N$ HSQC	upper	no	^{13}C or ^{15}N	yes	no	no
^{19}F, ^{31}P	none	no	^{19}F or ^{31}P	no	no	yes

2.1.1 Receptor-Based NMR Methods

The methods based on receptor-observation rely on the detection of chemical shift changes of the target resonances upon titration of the ligand. Such methods allow the identification of the ligand binding site on the receptor surface obtained by assigned protein NMR spectra, that require *a priori* knowledge of the protein three-dimensional structure (either from X-ray or NMR). Thus the identification of perturbations in assigned protein resonances in the presence of a test molecule, not only allows to verify if the compound is a receptor ligand, but also to localize its putative binding site. The localization of the binding sites also enables to immediately distinguish specific from non-specific binding. In addition, unlike ligand-based methods that rely on fast exchange, receptor-based methods permit the characterization of both higher and lower affinity interactions.

Nevertheless these NMR experiments are more demanding in terms of physicochemical properties of the receptor molecule. In fact, milligram quantities of soluble, non-aggregated target must be over-expressed and purified with a suitable isotope enrichment (e.g. ^{13}C, ^{15}N, and, sometimes, ^{2}H for proteins) essential for the resonance assignment of large (> 30 kDa) therapeutic targets. In addition, the sample has to be stable for the time required for sequential resonance assignment. All this process can be relatively long (some weeks) and represents the main limitation to the applicability of this approach.

2.1.2 Ligand-Based NMR Methods

Ligand-based methods observe and compare the NMR parameters of the ligand resonances in the presence and in the absence of the receptor. As they are based on the observation of the ligand resonances, they can be

exploited only for ligands with a low affinity (K_D values in the high nanomolar to the low millimolar range). Their use is mandatory when the NMR observation of the receptor is very difficult or prevented by either the inability to express and purify the target in heterologous systems with the desired labelling pattern, in the case of proteins the target dimensions (>100 kDa), or receptor poor solubility in physiological conditions that hamper its dissolution at higher concentrations.

In general, as previously introduced, these approaches are based on the exchange-mediated transfer of bound state information to the free state. This requirement implicates that the free and the bound state are in rapid exchange and that large ligand molar excess is adopted. Indeed, the process underlying most important NMR-based ligand binding screening experiments can be described by the following equation:

$$[E]+[L]\underset{K_{off}}{\overset{K_{on}}{\leftrightarrow}}[EL] \quad (1)$$

It represents a dynamic equilibrium involving three species: the free receptor E, the free ligand L, and the receptor-ligand complex, EL. The unimolecular rate constant k_{off} is inversely proportional to the mean lifetime τ_B of the receptorligand complex. The bimolecular rate constant, k_{on}, measures the probability of a productive encounter between the free receptor and ligand and is often assumed to be controlled by diffusion. k_{on} can vary between 10^7 and $10^9 M^{-1}s^{-1}$. The binding affinity can be described by the temperature-dependent equilibrium dissociation constant, K_D:

$$K_D = \frac{[E][L]}{[EL]} = \frac{K_{off}}{K_{on}} \quad (2)$$

Combining the definition of K_D with that of the bound receptor fraction:

$$P_B^E = \frac{[EL]}{[E]+[EL]} \quad (3)$$

These two equations can be combined giving the following:

$$P_B^E = \frac{[L]}{[L]+K_D} \quad (4)$$

where P_B^E is the fractional occupation of the receptor binding site by ligand L. When [L]« K_D, P_B^E is proportional to [L]. When [L] = K_D, the receptor is half saturated, that is, half of the receptor molecules exist in a one-to-one complex with the ligand. When [L]»K_D, the receptor is completely saturated and P_B^E =1.0; in this limit, every receptor binding site is occupied by one ligand, which on average exchanges with a second distinct ligand every ≈1/k_{off} seconds. According to these points, ligands with weak affinity have large K_D and thus require the addition of more ligand to saturate the receptor binding site. In the two-state equilibrium given by Eq. 1, ligand and receptor molecules will exist in either the free (L, E) or the complexed (EL) state. In the free state, both receptor and ligand retain their intrinsic NMR parameters. When ligand and receptor interact, the mutual binding affinity drives an exchange process that toggles both sets of molecules between the free and bound states. At the equilibrium, they adopt free and bound state populations ([E], [L], [EL]) consistent with Eq. 1. Under these conditions, the ligand transiently adopts the NMR parameters characteristic of the typically much larger receptor. Alternatively, from the receptor's perspective, the ligand transiently perturbs the binding site microenviroments, which may alter the distribution of conformations sampled by the ensemble of receptor molecules. In either case, the exchange modulates the NMR parameters of both molecules. It is the ability to experimentally distinguish these exchange-modulated parameters from those of the free state what enables NMR detection of the receptor-ligand interaction and, hence, NMR-based screening. The solution to the HMM equations (Bloch equation formalism by Hahn, Maxwell and McConnell) describes the behaviour of a

system magnetization on arbitrary exchange time scales. In NMR-based screening, however, these equations are almost never solved, and fast exchange is simply assumed. This assumption is made for two reason. The first is that the experimental conditions for ligand-based NMR screening are often well suited to fast exchange. These experiments are typically carried out with $L_T/E_T > 10$ (where L_T and E_T are total ligand and total receptor concentration respectively) and the binding compounds have $K_D \geq 100$ μM. If k_{on} is well approximated by diffusion-limited value (10^7 - 10^9 $M^{-1}s^{-1}$), then the slowest exchange rate K_{ex} values lie in the range $1000 < K_{ex} < 100000$ s^{-1}. Ligand based NMR screening methods are primarily ^1H based and, consequently, K_{ex} exceeds most differences in intrinsic ^1H relaxation rates and rotating frame precession frequencies, thus ensuring that the fast exchange assumption is valid. A second motivation for assuming fast exchange is the resulting algebraic simplicity. Generally, the NMR parameter Q becomes the simple average:

$$Q_{avg} = P_B Q_B + P_F Q_F \quad (5)$$

where P_B is the bound ligand fraction, and $P_F=1-P_B$. Here, Q_{avg} is the observed exchange-averaged parameter for the ligand (or the receptor) in the presence of the receptor (or the ligand). Observed differences between Q_{avg} and Q_F provide a signature of receptor binding and indicate a hit in a NMR screening based on that parameter. Q_{avg} is a simple population-weighted average and it can be applied to those parameters Q for which chemical shift modulations are not relevant. These parameters include longitudinal auto relaxation and cross-relaxation rates, rotating frame spin locking auto relaxation and crossrelaxation rates, and translational diffusion coefficients. The bound state contribution is P_BQ_B. The ability to detect binding with adequate sensitivity depends critically on P_BQ_B, being significant relative to P_FQ_F. However, typical screening conditions where

L_T»E_T make P_B «P_F. For these reasons, it is much preferred to measure NMR parameter Q that becomes amplified in the bound state.

A number of NMR variables can be monitored in ligand-based screening, such as (1) line broadening, (2) changes of the NOE effect (as in Saturation Transfer Difference (STD) spectroscopy,[5] Water-LOGSY,[6] transferred-NOESY (trNOESY)[7] and NOE pumping experiments[8]), (3) changes in the transverse and longitudinal relaxation rates of the ligand due to interaction with the receptor and/or spin labels[9], (4) changes in ligand diffusion coefficients measured by diffusion-edited NMR spectroscopy.[10]

These methods do not permit the identification of the ligand binding site, but, comparing the NMR parameters of a single molecule or of a mixture of compounds in the presence and absence of the receptor molecule, they do not require the production of milligram quantities of isotope labelled receptor and make the molecular weight of the target irrelevant; nay, most of these experiments become more sensitive with larger receptors and less than 1 mg of unlabelled protein is required for these experiments. In this way, new targets can be evaluated more rapidly and the experiments can be performed on a time scale useful for chemistry and high-throughput screens.

Of course, when receptor-based methods are applicable, they have to be chosen because of their potentially higher information. However, ligand-based screening is, in general, of broader applicability as a consequence of the limited number of low-molecular weight targets.

---

[5] Mayer M., Meyer B. (1999) Angew. Chem. Int. Ed. *38*, 1784–1788.
[6] Dalvit C. *et al.*(2000) *J. Biomol. NMR 18*, 65–68.
[7] Ni F. (1994) *Prog. NMR Spectrosc. 26*, 517–606.
[8] Chen A.D. *et al* (2000) J. Am. Chem. Soc. *122*, 414–415.
[9] Jahnke W. (2001) J. Am. Chem. Soc. *123*, 3149–50.
[10] Lin M. *et al.* (1997) J. Am. Chem. Soc. *119*, 5249–5250.

Below are described the most important NMR experiments that I performed during my thesis.

Saturation Transfer Difference (STD)

STD results from the difference of two experiments. In a first experiment, the "on-resonance" experiment, various receptor proton magnetizations are selectively saturated via a train of frequency-selective rf pulses. The rf train is applied to a frequency window that contains receptor resonances, but not resonances from the compounds. The saturation propagates from the point of application to other receptor protons *via* the vast network of intramolecular $^{1}H–^{1}H$ cross-relaxation pathways; this process of spin-diffusion is quite efficient, due to the typically large molecular weight of the receptor. Bound compounds pick up this saturation *via* inter-molecular $^{1}H–^{1}H$ cross-relaxation at the ligand–receptor interface. They then dissociate back into free solution where the saturated state persists, due to the small free state R_1 values ($R_1=1/T_1$). At the same time, more "fresh" ligand exchanges on and off the receptor while saturation energy continues to enter the system through the sustained application of rf. Thus, saturated free ligands accrue during the saturation time and this allows the "amplification" of the signal. For the on-resonance irradiation frequency values around -1 ppm are practical because no ligand nuclei resonances are found in this spectral region, whereas the significant line width of protein signals still allows selective saturation. If the ligand(s) show no resonance signals in the aromatic proton spectral region, the saturation frequency may also be placed there (7 ppm). In order to achieve the desired selectivity and to avoid side-band irradiation, shaped pulses are employed for the saturation of the protein signals.

A complementary reference experiment, the "off-resonance" experiment,

that applies the identical rf train for off-resonance, is then recorded; the spectrum yields a normal NMR spectrum of the ligand-protein mixture.

The “on-resonance” and “off-resonance” experiments are recorded in an interleaved manner and subtracted. A difference spectrum is produced in which only the signals of the ligand that have felt the transfer of saturation from the protein do appear. The receptor resonances will not be visible, or will be scarcely visible, as consequence of their minimal concentration or because of R_2 relaxation-filtering (spin-lock) applied just prior to detection (Fig. **1**).

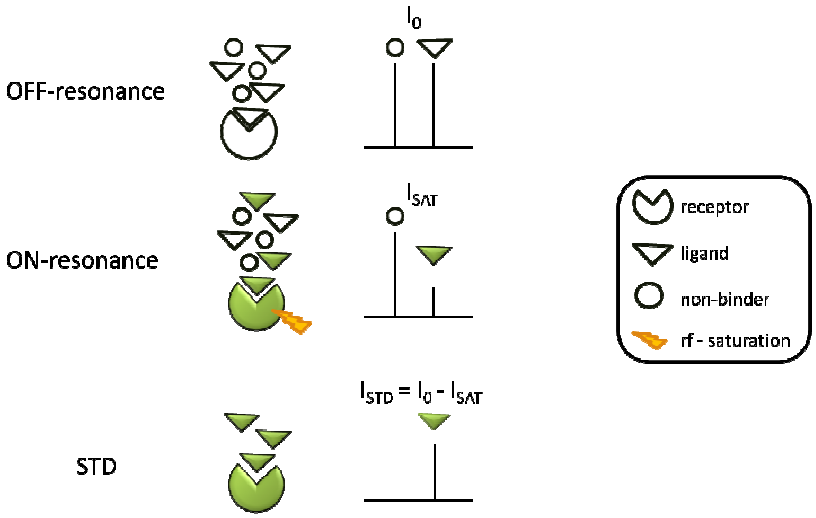


Figure 1. STD spectrum is generated by subtraction of “on-resonance” spectrum from “off-resonance” spectrum.

If binding is very tight, and consequently off-rates are in the range of 0.1–0.01 Hz, the saturation transfer to ligand molecules is not very efficient. This is usually the case for K_D values less than 1 nM. If the K_D values are 100 nM, or larger, fast exchange of free and bound ligands leads to a very efficient build up of saturation of the ligand molecules in solution. On the other hand, when binding is very weak, the probability of the ligand being in the receptor site becomes very low, which results in weak STD signals.

Thus, STD can be used from tight binding up to a K_D of about 10 mM.

The intensity of the STD signals depends, among other things, on the irradiation time / saturation time and on the excess of ligand molecules used.[4] The more ligand is used and the longer the irradiation time, the stronger the STD signal is. In general, the irradiation time is 2s and until a 100-fold excess of ligand is used to give good results. In fact, upon dissociation, the saturation of the ligand is transported in solution, where it accumulates as a result of the slow decline of the saturation by relaxation processes in solution. Before ligands in solution have lost their saturation, the process of association (followed by dissociation) can occur many times and thus put many more saturated ligands into solution. The maximum net effect of saturation on ligand protons occurs if a large excess of ligand is used, because it is very unlikely that a ligand with saturation re-enters the binding site. From the high ligand:receptor ratios it is clear that the amount of receptor required for the measurements is very small. The ligand saturation is higher for those protons that are in closer contact with the receptor. For those ligand protons that interact with protein protons through an intermolecular NOE, a decrease in intensity is observed in the "on-resonance" spectrum, while no intensity variations take place in the "off-resonance" spectrum. As already stated, subtraction of the "on-resonance" from the "off-resonance" spectra leads to a difference spectrum, in which only the signals of the protons that were attenuated by saturation transfer are visible (Fig. **1**). In this way, STD spectroscopy not only furnishes evidences of ligand-protein binding, but also allows to obtain the ligand epitope mapping, that is the identification of the ligand regions involved in the interaction with the receptor. This kind of information is of primarily importance for pharmacophore group identification in the rational drug design process.

Solvent Saturation Transfer

A variant of STD exists in which saturation transfer involves bound water instead of receptor. It relies on the selective saturation (or inversion) of water resonance and the following transfer of magnetization to the ligand. Water molecules are often found at the receptor-ligand interface forming a ternary complex. The magnetization transfer from bulk water to ligand takes place *via* labile receptor protons (exchangeable protons) within but also distant from the ligand-binding pocket, as well as from water molecules that reside for long time within the binding pocket. In the presence of small ligands, the intermolecular water-ligand NOE becomes negative only if the ligand is bound to the receptor. This is a consequence of the fact that the correlation time is much longer for the ternary complex water-ligand-receptor than for the water-ligand complex. So, if the residence time of water molecules is long enough, negative NOE effects are observed between water and ligand protons. In this experiment, the resonances of non-binding compounds appear with opposite sign and tend to be weaker than those of the interacting ligands. The main technique exploiting this principle is the Water-Ligand Observed via Gradient SpectroscopY (WaterLOGSY).[6]

This experiment takes advantage from the large reservoir of bulk water magnetization, that makes this technique very sensitive and allows it to be used with protein concentrations as low as 100 nanomolar. The practical aspects and range of applicability of the WaterLOGSY experiment are analysed in detail in Dalvit *et al*.[11]

[11] Dalvit C. *et al.*(2001) Journal of Biomolecular NMR *21*, 349–359.

Transferred NOESY (trNOESY)

NOE effects (NOEs) have been broadly exploited in determining the 3D structure of molecules in solution,[12] recently they have found important applications also studying the field of ligand-protein interaction. When ligand molecules bind to receptor proteins, the NOEs undergo drastic changes leading to the observation of transferred NOEs (trNOEs). The observation of trNOEs relies on the existence of different tumbling times τ_c for the free and bound molecules. Low or medium molecular weight molecules (MW < 1000 Da) have short correlation times τ_c and, as a consequence, such molecules exhibit positive NOEs. Large molecules, however, exhibit strongly negative NOEs. When a small molecule (ligand) is bound to a large molecular weight protein (the protein receptor molecule), it behaves as a part of the large molecule and adopts the corresponding NOE behaviour, that is, it shows strong negative NOEs, so-called trNOEs. These trNOEs reflect the bound conformation of the ligand. Binding of a ligand to a receptor protein can thus easily be distinguished by looking at the sign and size of the observed NOEs (Fig. **2**).

[12] Neuhaus D. *et al.*(2000) published by Wiley.

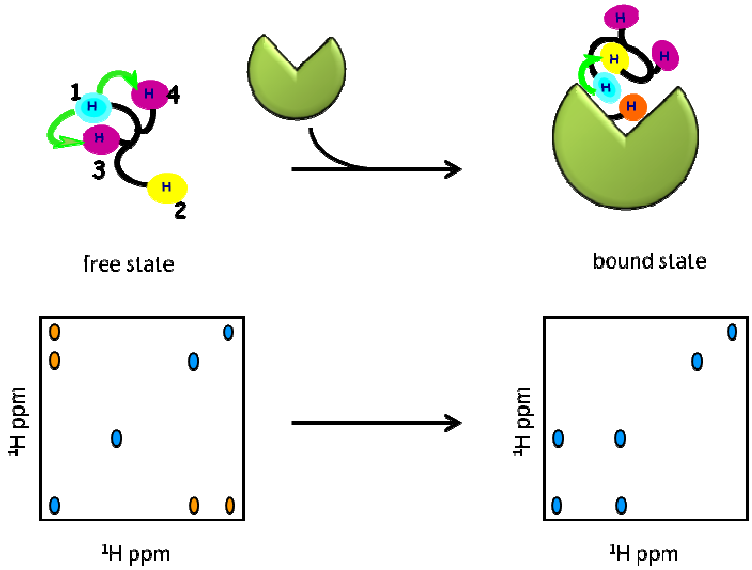


Figure 2. Left. Schematic representation of a NOESY spectrum for a free ligand. Cross peaks and diagonal peaks have different sign. Right. Schematic representation of a trNOESY spectrum recorded for an exchanging ligand–protein system. Cross peaks and diagonal peaks have the same sign, as expected for a large molecule, thus indicating binding to the protein.

However, one of the major drawbacks of this experiment is the possible existence of spin diffusion effects, which are typical for large molecules. In this case, apart from direct enhancements between protons close in space, other spins (including those of the receptor) may mediate the exchange of magnetization, thus producing negative cross peaks between protons far apart in the macromolecule. Thus, protein-mediated, indirect trNOE effects may lead to interpretation errors in the analysis of the ligand bound conformation. However, using trNOEs in the rotating frame (trROESY) experiments[13,14] it is possible to distinguish which cross peaks are dominated by an indirect effect, usually mediated by a protein proton, and

[13] Arepalli S.R. *et al.* (1995) J. Magn. Reson. Ser. B *106*, 195-198.
[14] Asensio J.L. *et al.* (1995) Eur. J. Biochem. *230 (2)*, 621-33.

hence distinguish direct from indirect enhancements.

The estimated range of binding affinities that can be probed by trNOESY is $100nM \leq K_D \leq 1mM$.[15]

Thus, the trNOE method allows fast screening of putative binders respect to a specific target receptor and, at the same time, permits the knowledge of the recognized conformation of the ligand bound to the receptor, with considerable implications for a rational structure-based drug design.

Competitive Experiments

The main problem with most ligand-detected experiments is the limited affinity range (K_D within the range 10^{-3}–10^{-7} M) they allow to investigate. In order to avoid this problem and to compare the relative binding affinity of different compounds, titrations with competitive STD or WaterLOGSY can be performed.

To perform this kind of experiments, the tested ligands present in the mixture should have at least one non-overlapping signal in the spectrum. The reporter ligand is then held at a constant concentration in the presence of the receptor, whereas a second ligand is titrated into the solution, by increasing its concentration. For a quantitative analysis, at each addition point, it is necessary to acquire a STD or a WaterLOGSY spectrum and to measure the decrease in the intensity of the signals of the first ligand. Conversely, for a qualitative information, it is enough to record only a spectrum in which the two ligands are present at the same concentration. Indeed, it is possible to assess whether the second compound has a higher binding affinity, which enables it to compete with the first molecule.

[15] Mayer M. and Meyer B. (2000) J. Med. Chem. *43(11)*, 2093-2099.

As already stated, in the case in which a ligand has a high binding affinity (K_D « 1μM), it will be subjected to a slow chemical exchange compared to the NMR-time scale and it will not be identified with the normal experiments. However, the competitive experiments allow to overcome this limitation, since, indirectly, the decrease and eventual disappearance of STD signals of the first ligand, (if the mixture is kept at constant concentration) is the index of competing ligand with high affinity, even if the latter has not observable signals.

The actual observation of a competitive binding then indicates that the different molecules bind to the same binding site and allows to derive important information on the correlation between the structure and the binding affinity.[16]

3. NMR As A Tool To Follow Enzymatic Reactions

Kinetic Measurements

NMR spectroscopy has become a valuable tool for the study of enzyme kinetics in real time. In particular, in order to follow these phenomena, it can be used to acquire 1D ^{1}H NMR spectra over a certain time period.

The great advantage of this method is that this is a non-destructive technique, which allows to follow directly the reaction kinetics, without the need to couple secondary reactions of derivatization, monitoring either the progressive decrease intensity of the signals of the substrates, or the consequent buildup of the reaction products. The only essential requirement for investigating enzyme-mediated reactions by NMR

[16] Blume A. *et al.* (2006) Journal of biological chemistry *281(43)*, 32728-32740.

spectroscopy is that substrates and products must show at least one resolved signal, in order to follow its evolution over time.[17]

4. High Resolution Magic Angle Spinning-NMR

High-Resolution Magic-Angle-Spinning (HR-MAS) NMR spectroscopy is a well established tool for studying heterogeneous system, especially semisolid materials such as lipid membranes, drug delivery system, cell suspensions, biopsy samples, molecules adsorbed on a solid support, and resin-bound molecules. In such system, MAS sufficiently suppresses the disturbing anisotropic interactions, such as magnetic susceptibility, chemical shift anisotropy or even dipolar coupling. At modest rotation frequencies, well-resolved ^{1}H-NMR spectra can be observed, and the straightforward application of NMR methods commonly used for liquid sample is possible.[18] Indeed, the line width of a NMR resonance depends strongly on the microscopic environment of the nucleus under study. Interactions such as the chemical shift and dipole-dipole coupling between neighboring spins are anisotropic and impose dependence on the NMR frequency based on the orientation of a spin or molecule with respect to the main magnetic field direction. Furthermore, the magnetic susceptibility of the sample and susceptibility differences within the sample lead to a broadening of the resonances.[19] In liquid state spectroscopy the rapid isotropic motion of the molecules averages the anisotropic interactions, resulting in an isotropic chemical shift frequency and a disappearance of the line broadening due to dipolar couplings. Furthermore the sample geometry, a cylinder parallel to the main magnetic field, is chosen such that

[17] Haselhorst T. *et al.* (2005) Biochemical and biophysical research communications *327*, 565-570.
[18] Pampel A. and Engelke F. Spin Report, www.bruker.it
[19] Bruker-Biospin HR-MAS NMR Tutorial.

the susceptibility broadening is minimized. In solid samples, on the other hand, the lack of molecular mobility results in broad lines. This line broadening can be reduced by spinning the sample rapidly around an axis which is oriented at an angle of 54.7° with the direction of the magnetic field (Fig. **3**). By spinning at this so-called 'Magic Angle', at a rate larger than the anisotropic interactions, these interactions are averaged to their isotropic value, resulting in substantial line narrowing. In fact, MAS removes the magnetic susceptibility broadening.

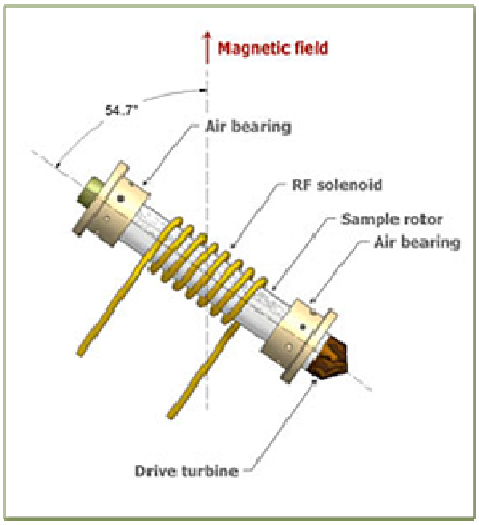


Figure 3. Magic angle spinning involves rapid rotation of a sample inclined at 54.7° to the static field.

In addition to pure solids or pure liquids, there is a wide range of materials which exhibit either reduced or anisotropic mobility. For example, HR-MAS is a very useful technique for detection, identification and quantification of compounds bound to a polymer support. Spinning of the sample at the magic angle, when combined with the swelling of the resin, usually results

in high resolution liquid spectra like ^{1}H and ^{13}C for compounds still attached to the resin[20]. (Fig. **4**)

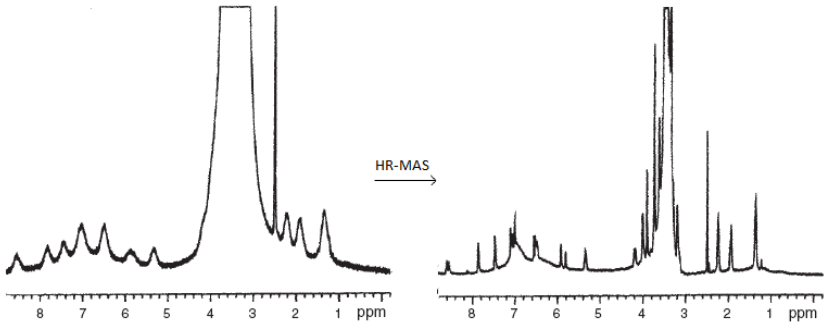


Figure 4. The conventional ^{1}H NMR spectrum of a resin bound analyte (on the left) and the ^{1}H HR-MAS spectrum of the same sample.[21]

[20] Ramadhar T.R. *et al.* (2007) Magn. Reson. Chem. *46*, 30-35.
[21] Claridge T. D. W. Tetrahedon Organic chemistry *27*, Elsevier.

5. Overview On The Studied Targets

Hereinafter I introduce the different ligand-macromolecule pairs of biological and biomedical relevance that have been studied exploiting NMR techniques.

5.1 A Target For Antibiotic Design: Arabinose 5-Phosphate Isomerase

Since the discovery of the antibiotic properties of penicillin in 1929, a great number of antibacterial compounds have been produced, contributing to an improvement of human health and medicine.[22] These years of antibiotic use, and more often of abuse, has led to the spread of resistant bacteria and superbugs, leading to an increase of mortality rates caused by these infectious agents.[23,24] As these pathogens cause diseases in all body systems, the consequences of their re-emergence have a great impact both economically and socially.[23] Thus, there is an urgent demand for the development of new molecular entities targeting novel essential bacterial enzymes.[24]

In order to achieve this goal, it is therefore necessary to dissect biochemical pathways critical for pathogens survival and virulence and to evaluate them as target for antibiotic development. In this context, it has long been recognized that, among microbial components, lipopolysaccharide (LPS) is an essential moiety in the outer membrane of Gram-negative bacteria. These bacteria indeed are typically surrounded by two membranes: an inner (cytoplasmic) membrane (IM) and an outer membrane (OM) separated by an aqueous compartment, the periplasm, containing a layer of

[22] Nie W. *et al*. (2012) Clinical Science *122*, 575-580.
[23] Kumar S. *et al.* (2012) Int. J. Mol. Sci. *13*, 4484-4495.
[24] Bhagunde P. *et al.* (2012) Antimicrobial Agents and Chemotherapy *56*, 5, 2237-2240.

peptidoglycan. Whereas the IM is a symmetric bilayer composed of phospholipids, the OM is an asymmetric bilayer composed of phospholipids in the inner and lipopolysaccharides (LPS) in the outer leaflet, respectively.

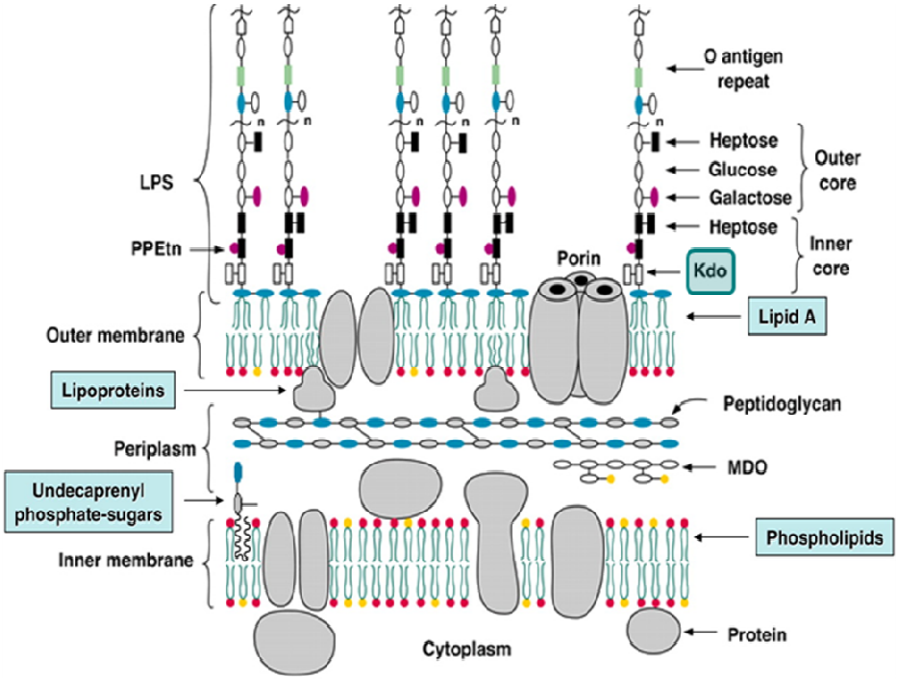


Figure 5. Structure of the Gram-negative envelope.

The peculiar permeability barrier properties that distinguish the OM from the IM are mainly due to the presence of LPS in its outer leaflet, which provides resistance to several external noxious agents, thus allowing this group of bacteria to inhabit hostile environments.[25]

Within the host, LPS (also known as endotoxin) is recognized by CD14 and the Toll-like receptor TLR4-MD2 complex. CD14 binds and concentrates LPS, presenting it to TLR4-MD2 that triggers the biosynthesis of diverse

[25] Nikaido H. (2003) Microbiol. Mol. Biol. Rev. *67*, 593.

mediators of inflammation, such as TNF-α, IL-1 and IL-6, which activate the production of costimulatory molecules for adaptive immunity.[26,27,28]

Lipid A is the LPS endotoxin toxic portion and is the crucial moiety in the interaction with TLR4-MD2. Excessive response to LPS, caused by blood-circulating microbial antigens in affected patients, results in severe sepsis, a rapidly progressing inflammatory disease with up to 29% mortality.

Structurally, LPS is a complex glycolipid that consists of three covalently linked domains: a hydrophobic domain termed lipid A (a glucosamine-based phospholipid), a non-repeated oligosaccharide core and a distal polysaccharide of various lengths (the so-called *O*-antigen). The core oligosaccharide provides the link between lipid A and the structurally diverse *O*-antigen polysaccharide chain and can be further divided into the inner and outer core. The structure of the outer core is somewhat variable[26], whereas the inner core region is more conserved, with one structural element 3-deoxy-D-manno-oct-2-ulosonic acid (Kdo) always present in all inner core so far analyzed.[29] The minimal LPS structure required for growth of *E. coli* is composed of two Kdo residues attached to lipid A (Kdo2-lipid A or Re LPS), suggesting the importance of this complex in order to maintain structural integrity and cell viability. The ubiquitous nature of Kdo within LPS structures has prompted detailed studies on its biosynthesis. The pathway starts with the enzyme D-arabinose 5 phosphate (A5P) isomerase (API), which converts the pentose pathway intermediate D-ribulose 5-phosphate into A5P. Subsequently, A5P is condensed with phosphoenolpyruvate to form Kdo 8-phosphate (Kdo8P) (KdsA), hydrolyzed to Kdo (KdsC), activated as the sugar nucleotide CMP-Kdo (KdsB), before

[26] Raetz C.R. and Whitfield C. (2002) Annu. Rev. Biochem. *71*, 635.
[27] Miyake K. (2004) Trends Microbiol. *12*, 186.
[28] Munford R.S and Varley A.W. (2006) PLoS Pathog. *2*, e67.
[29] Holst O. (2007) FEMS Microbiol Lett. *271*, 3.

being finally transferred from CMP-Kdo to the acceptor lipid IVA (WaaA) (Fig. **6**). The late acyltransferases LpxL and LpxM transfer the fatty acids laurate and myristate, respectively, to Kdo2-lipid IVA to form the characteristic acyloxyacyl units of hexa-acylated Kdo2-lipid A.

The inhibition of the Kdo biosynthesis leads to the arrest of cell growth and API represents a logical control point, because it catalyzes the first step in the Kdo pathway. The high homology level within Gram-negative bacteria along with its favorable localization and the fact that it lacks a human counterpart make API an attractive target for the development of potential antibacterial agents.

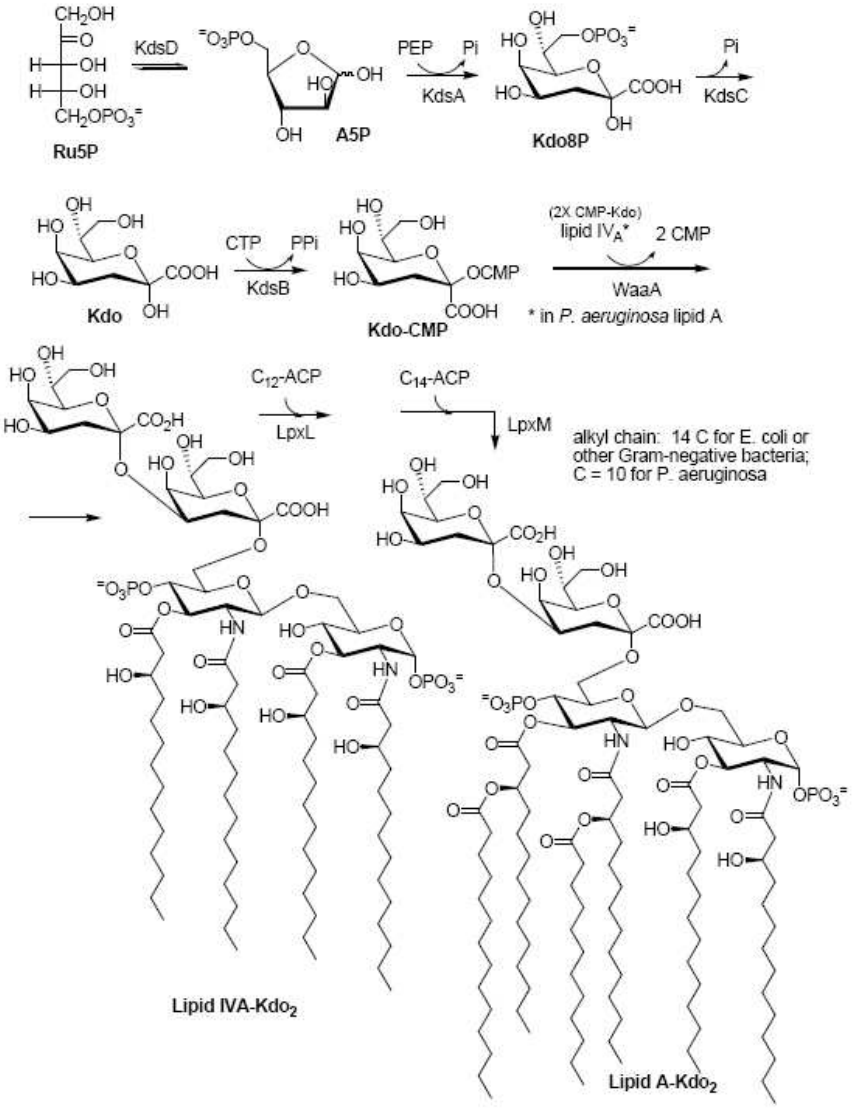


Figure 6. Kdo biosynthetic pathway and its incorporation to Lipid IVA.

In *E. coli* K-12, there are two API genes (*kdsD* and *gutQ*) which have been cloned, characterized, and shown to have nearly identical biochemical properties. While virtually all sequenced genomes of Gram-negative bacteria encode KdsD, only a subset of *Enterobacteriaceae* also encode

GutQ. The biological significance of this apparent API redundancy in *E. coli* is presently unknown, though it is has been shown that either gene alone can support LPS synthesis at near wildtype levels.[30] In addition to these two enzymes, there is a third paralog gene found in pathogenic strains of *E. coli*, as CFT073, K1 and K5, coding for KpsF and involved in the biogenesis of the polysialic acid capsule.

KdsD is the prototype of the sugar-isomerase, while *gutQ* and *kpsf* are paralogous genes, resulting from successive duplications of *kdsD* and associated to other specific pathways, but able (at least for *gutQ*) to replace *kdsD*.

The gene *kdsD* encodes for a protein of 328 amino acids; the molecular mass of each subunit, as determined by electrophoresis SDS-PAGE under reducing conditions is about 35 kDa, while the molecular mass of the native protein is 122 ± 5 kDa, indicating the tetrameric nature of API.

KdsD presents the N-terminal domain of 210 amino acids, called SIS (Sugar Isomerase), commonly found in the aldo-keto isomerases and proteins that regulate the expression of genes involved in the synthesis of phosphosugars.[31] The multiple alignment of the SIS domains belonging to different proteins shows that there are several conserved regions, that are involved in non-specific binding of the sugar-phosphate group, catalysis and in structural integrity of the domain, while the more variable regions may determine the substrate specificity. The structure of the SIS domain can be described as a 3-layer α-β-α sandwich, an architecture typical of the Rossmann fold, which is characterized by the presence of an open twisted parallel β-sheet surrounded by α-helices on both sides. However, proteins

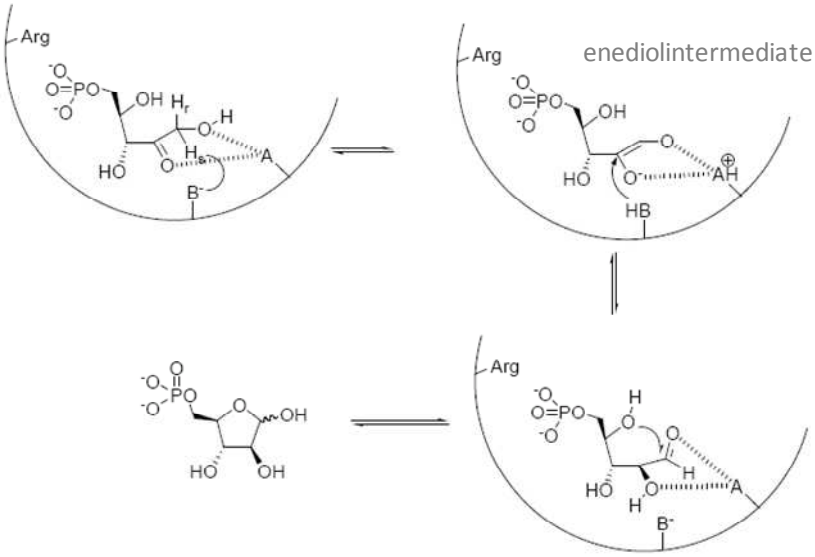
[30] Meredith T.C. *et al.* (2006) ACS chemical biology *1* (1), 33-42.
[31] Meredith T.C. and Woodard R.W. (2003) The Journal of Biological Chemistry *278* (35), 32771-32777.

containing the SIS domain usually do not share a significant sequence similarity and are characterized by a variety of protein architectures.[32]

KdsD presents in its C-terminal region two CBS domains (Cystathionine β-synthetase), each consisting of 50-60 amino acids; these motifs have been identified in many proteins and, although their exact function is still not known, it has been proposed that they play a regulatory role and are involved in various disease states.[31]

The overall structure of *E.coli* KdsD SIS domain has been recently predicted by homology modeling;[32] however, structural and catalytic details are still scarce. Each KdsD monomer presents a Rossmann fold and predicted residues involved in the catalytic mechanism have estimated pKa values of 6.55 and 10.34[31] suggesting the occurrence of a histidine or possibly a carboxylate along with a lysine or arginine at the active site. Similar amino acids are known to be involved in the catalysis of other sugar isomerases, suggesting a common mechanism. A plausible mechanism was formulated and involves a *cis*-enediolate intermediate as shown in Scheme **1**. A basic group on the enzyme probably removes the pro-*S* proton on C-1 of ribulose 5-phosphate and re-protonates the substrate at C-2 on the same (*si*) face of the enediol.

[32] Sommaruga S. *et al.* (2009) Biochemical and Biophysical Research Communications *388*, 222–227



Scheme 1. Proposed KdsD mechanism.

Homology modelling studies on KdsD suggest that Lys59, Glu111 and Glu152 are confined close to the presumed active site.(Fig. **7**)

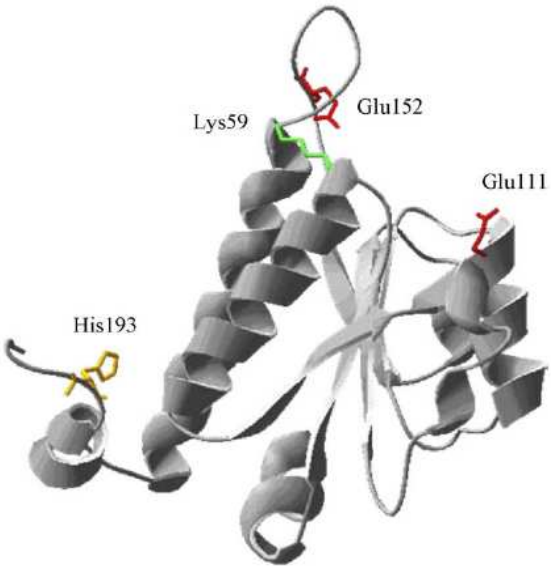


Figure 7. 3D model of *E.coli* KdsD SIS domain. It was modeled on the template structure PDB ID: 3etn. The predicted catalytic residues are indicated in sticks.

As highlighted in Fig.**7** , also His 193 was selected as a residue that can play a role in the catalysis. Indeed, even if His193 is localized at the C-terminal of the KdsD SIS domain and seems far from the putative active site, it can be contributed by another monomer of the tetramer. From this analysis the authors also evidenced that a common trait of all the sugar isomerases studied is the presence of a cluster of Ser and Thr residues near the active site, which are most probably involved in binding the phosphate moiety of phosphosugars.[32] In order to better characterize the involvement of the selected residues in the enzymatic catalysis, they were mutated to alanine: the Lys59Ala and His193Ala mutants showed a drastic loss of activity, indicating these two residues are functionally relevant.[33]

Subsequently also a crystal structure of a mutant SIS domain of KdsD (KdsD Lys59Ala) was solved at 2.6 Å resolution.[34] It seems to suggest that the proposed active site is formed by contributing residues from two chains, (chains A/C, and B/D). Structure-based sequence alignments have highlighted a highly conserved His residue (His88; Fig.**8**) not previously considered, as a possible residue relevant for catalysis.

[33] Cipolla L. *et al.* (2010) Nat. Prod.Rep. *27*, 1618–1629.
[34] Gourlay L.J., Sommaruga S. *et al.* (2010) Protein Science *19*, 2430-2439

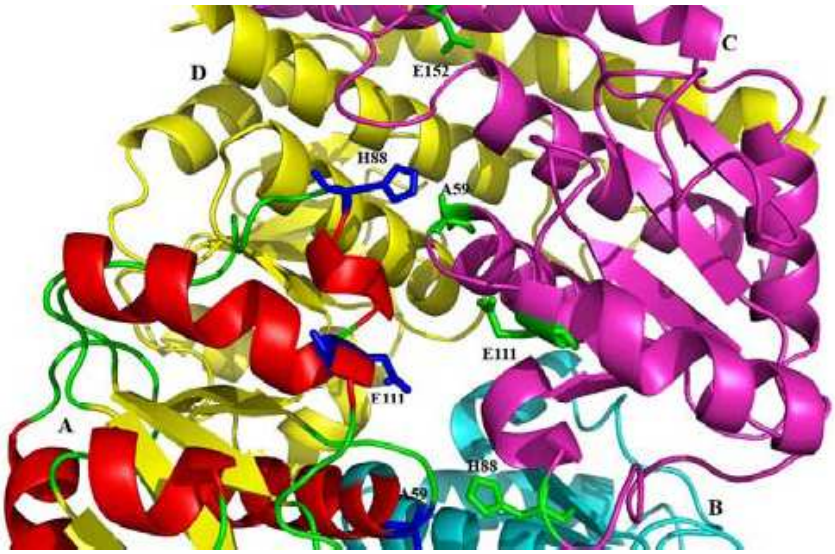


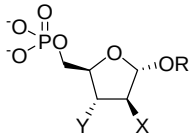
Figure 8. Crystal structure of KdsD Lys59Ala SIS domain. The four different SIS monomer are indicated with capitals.

It has been shown that *E.coli* KdsD follows a standard Michaelis-Menten kinetic and the kinetic parameters were determined for both ribulose-5-phosphate (K_M = 0.35 ± 0.08 mM; apparent K_{cat} = 255 ± 16 s^{-1}) and arabinose-5-phosphate (K_M = 0.61 ± 0.06 mM; apparent K_{cat} is 157 ± 4 s^{-1}). The reaction equilibrium at 37 °C, is shifted towards the formation of A5P (68 ± 2% versus 32.2 ± 2%) and the optimum pH is 8.4 (measured in 1,3-bis(tris(hydroxymethyl)methylamino)propane (BTP) buffer, with 1mM Ethylenediaminetetraacetic acid (EDTA)). It was demonstrated that recombinant *E.coli* KdsD contained approximately one zinc atom per monomer after purification. The metal was not required for KdsD activity, as it could be removed by treatment with EDTA followed by gel filtration and this resulted in nearly a 2-fold increase in observed activity. In addition, the activity was increased upon the *in situ* addition of excess EDTA. It was also shown that KdsD is extremely sensitive to d^{10} metal inhibition, with an estimated IC_{50} of 1–3 µM for zinc. It was speculated that the mode of

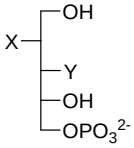
inhibition may be through coordination to active site carboxylate and/or histidine residues, which are commonly encountered in high affinity zinc ligands and that metal ion inhibition may be a mechanism to control activity *in vivo*, although the biological significance in terms of API is unknown. The enzyme seems to have a high substrate specificity as neither sugars with a shorter carbon chain, D/L-glyceraldehyde and D-erythrose-4 phosphate, nor the pentoses and hexoses phosphorylated tested (D-ribose-5 phosphate, D-glucose-6 phosphate, D-glucose-1 phosphate, D-glucosamine-6 phosphate and D-mannose-6 phosphate) are converted to their respective ketoses, within the limits of the colorimetric assay.[31]

In 1984 a series of compounds (Fig. **9**), which are analogues of the substrates or intermediate, were tested by Bigham and co-workers[35] as inhibitors of A5P isomerase with the belief that a good inhibitor would stop bacterial growth or render the cells more susceptible to other antibiotics or natural defenses.

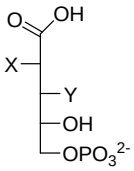
[35] Bigham E.C. *et al.* (1984) J.Med.Chem. *27*, 717-726.



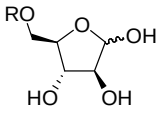
1: R = H; X = OH; Y = H (IC_{50} 66.7mM)
2: R = H; X = F; Y = OH
3: R = Me; X = OH; Y = F (IC_{50} 91.3 mM)
4: R = H; X = OH; Y = F (IC_{50} 61.3 mM)
5: R = H; X = OH; Y = OMe
6: R = H; X = OH; Y = OCH_2Ph
7: R = H; X = αOH[a]; Y = OH (IC_{50} 80.0 mM)
8: R =H; X = H; Y = OH



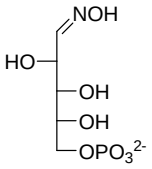
15: X = H; Y = OH
16: X = αOH[a]; Y = OH (IC_{50} 20.0 mM)
17: X = OH; X = OH (IC_{50} 0.7mM)
18: X = F; Y = OH (IC_{50} 5.3 mM)
19: X = OH; Y = F (IC_{50} 100.0 mM)
20: X = OH; Y = OBn (IC_{50} 233 mM)



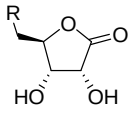
9: X = OH; Y = OH (IC_{50} 0.53mM)
10: X = αOH[a]; Y = OH (IC_{50} 1.6 mM)
11: X = F; Y = OH (IC_{50} 4.7 mM)
12: X = OH; Y = F (IC_{50} 62.0 mM)
13: X = H; Y = OH



21: R = SO_2NH_2 (IC_{50} 19.2 mM)
22: R = $C(O)NHSO_3Na$ (IC_{50} 19.2 mM)
23: R = NO_2
24: R = $C(O)CH_2COMe$
25: R = $SO_2(4\text{-}NH_2Ph)$



14 (IC_{50} 8.7 mM)



26: R = Cl
27: R = OH (IC_{50} 26.9 mM)
28: R = $MeCOCH_2CO_2^-$
29: R = $NH_2SO_3^-$(IC_{50} 8.4 mM)
30: R = $MeSO_3$ (IC_{50} 35.9 mM)
31: R = $NaOSO_3^-$ (IC_{50} 8.4 mM)
32: R = $4\text{-}NH_2PhSO_3^-$
33: R =$OCOCO_2^-Na^+$ (IC_{50} 12.0 mM)

Figure 9. Structure of proposed inhibitors for API. [a] α implies a *ribo* configuration. IC_{50} is reported only for active compounds. (Compounds: 1-8: A5P analogues; 9-13:aldonic acid phosphates; 15-20: alditol phosphates; 21-25: arabinose derivatives; 26-33: ribonolactones).

From all these activity tests, performed by using a discontinuous cysteine-carbazole colorimetric method[36] a few conclusions could be drawn. In the series of phosphorylated sugars, the order of isomerase inhibitory activity was as follows: aldonic acids > alditols > aldoses, reflecting their ability to

[36] Dische A., Borenfreund E. (1951) J. Biol. Chem. *192*, 583.

mimic the proposed enediol intermediate. Nonphosphorylated sugars were found to be much less active. The best inhibitor was the four-carbon aldonic acid erythronate-4 phosphate (IC_{50} = 0.22mM), however this compound is also able to inhibit the ribulose-5 phospate isomerase, as reported by Wolfenden[37] and this is undesirable since ribulose 5-phosphate isomerase is also present in the human host. None of the compounds displayed antibacterial activity *in vitro*, probably because they do not penetrate the bacterial cell.[35]

Very recently, Woodard and colleagues synthesized other few putative inhibitors of *Francisella tularensis* KdsD (Fig. **10**).[38] They selected hydroxamates enediol mimics as this class of compounds were reported to be active in similar systems. In fact the synthesized compounds were already known as 6-phosphogluconate dehydrogenase (6PGDH) of *Trypanosoma brucei*.[38]

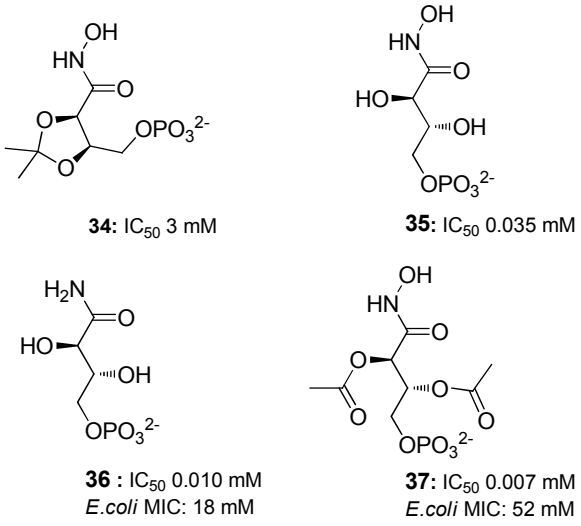


Figure 10. Structures and inhibitory activity of the compounds synthesized by Woodard and coworkers.

[37] Wolfenden R. and Woodruff W. (1979) J. Biol. Chem. *254*, 5866.
[38] Yep A. *et al.* (2011) Bioorganic and Medicinal Chemistry Letters *21*, 2679-2682.

Performing the same colorimetric assay already used by Bigham and colleagues,[35] it was found that these compounds seem to inhibit KdsD, even if their activity is lower compared to their potency against 6PGDH; nevertheless, the most potent KdsD inhibitior reported to date was discovered (compound **37**).

The rationale behind the inhibition properties of the reported molecules is not immediately clear, as all these results were obtained with a colorimetric assay that does not give any further indication about the binding mode of the compounds and does not allow to verify if the tested ligand are also KdsD substrates.

More efforts are therefore needed to understand KdsD binding properties and to develop antibiotics targeting this interesting enzyme.

5.2 Prion Protein (PrP) And Prion Diseases

Prion diseases are transmissible, progressive and invariably fatal neurodegenerative disorders associated with misfolding and aggregation of a host-normally-encoded cellular prion protein, PrPC. They occurr in a wide range of mammalian species, including human. Human prion diseases can arise sporadically, can be hereditary or acquired. Among the sporadic human prion diseases, Cruetzfeldt-Jacob disease (CJD), fatal insomnia and variably proteasesensitive prionopathy are found; genetic or familial prion diseases are caused by autosomal dominantly inherited mutations in the PrPC encoding gene and include genetic or familial CJD, fatal familial insomnia and Gerstmann-Sträussler-Scheinker syndrome. Kuru, iatrogenic CJD and a new variant form of CJD that was transmitted to humans from affected cattle via meat consumption belong to the acquired human prion diseases and they account for only 5% of cases of human prion diseases.[39]

The key features of nearly all prion diseases are transmissibility and a foamy pattern (spongiosis) of brain tissue that is visible on histological sections and that is due primarily to intraneuronal vacuoles containing membrane fragments and, sometimes, degenerating organelles.[40]

The term “prion” was coined in 1982 by Prusiner to describe the "proteinaceous infectious particles" causing scrapie, the prototypic prion disease, in sheep.[41,42,43] He suggested that this pathology and a collection of other neurodegenerative diseases were all due to a common process: a misfolded protein that propagates and kills brain cells. Indeed, this protein-only hypothesis asserts that the transmission of the pathology does not

[39] Mahmood I. (2011) Virology Journal *8*, 559.
[40] Aguzzi A. and Steele A.D. (2009) Cell *137*, 994-996.
[41] Prusiner S.B. and DeArmond S.J. (1990) Monogr Pathol. *32*, 86-122.
[42] Prusiner S.B. (1998) Proc. Natl. Acad. Sci. USA *95*, 13363–13383.
[43] Prusiner S.B. (1982) SCIENCE *216*, 136-144.

require nucleic acids, and that the PrP abnormal conformer (PrPSc from scrapie prion protein) represents a self-perpetuating infectious agent. Once heretical, this model is now gaining rapid acceptance, especially in view of recent advances in generating infectious PrPSc *in vitro*.[44] Moreover it was observed that PrP-knockout mice do not propagate PrPSc infectivity and are fully resistant to prion diseases.[45]

The gene encoding PrP is *PRNP* and is located in the short arm of human chromosome 20 and in a homologous region in mouse chromosome 2; the entire ORF of all known mammalian *PRNP* genes is contained in a single exon, eliminating the possibility that the misfolded PrP (PrPSc) arises from alternative RNA splicing, although mechanisms such as RNA editing, protein splicing and alternative initiation of translation remain a possibility.[46] The expression pattern of PrPC is broad, developmentally regulated and includes the nervous system (with high PrPC levels in synaptic membranes of neurons and on astrocytes), secondary lymphoid organs, skeletal muscle, kidney and heart. Moreover PrPC is highly conserved among mammals, and paralogues.[47] The normal cellular function of PrPC has remained elusive and the proposed physiological cellular roles of PrPC are many, ranging from copper internalization and homeostasis to antiapoptotic activities, protection against oxidative stress, cell adhesion, cell signaling, and the modulation of synaptic structure and function, but none of these functions have been convincingly demonstrated.[48] Studies with PrP knockout mice have not been very useful to determine the function of PrPC, because the removal of the protein produces mild phenotypes. However, the subtle abnormalities in these animals include alterations in neuronal transmission

[44] Cobb N.J. (2007) PNAS *104*, 48, 18946–18951.
[45] Biasini E. and Harris D.A.(2012) Future medicinal chemistry *4* (13), 1655-1658.
[46] Aguzzi A. and Calella A.M. (2009) Physiol. rev. *89*, 1105-1152.
[47] Nuvolone M. *et al.* (2009) FEBS Letters *583*, 2674–2684.
[48] Cobb N.J. and Surewicz W.K. (2009) Biochemistry *48,* 2574–2585.

and electrical activity, defective neurogenesis, alterations in circadian rhythm and increased sensitivity to hypoxia, ischemia and seizures. Moreover, ablation of PrPC triggers a chronic demyelinating polyneuropathy in mice, indicating that neuronal PrPC expression might be essential for myelin maintenance.[49] The list of putative PrPC binding partners is equally long; some of these cellular cofactors have been suggested to contribute not only to normal PrPC function but also to the conformational conversion process, leading to the formation of PrPSc.[48]

Following the cleavage of the signal peptide, mature PrPC mainly localizes to detergent-resistant subdomains known as "lipid rafts" on the outer surface of the plasma membrane via a C-terminal glycophosphatidylinositol (GPI) anchor.[48] PrPC is a ≈210 amino acid protein containing a N-terminal random coil sequence of ≈100 amino acids (from amino acid 23 to 124) and a C-terminal globular domain, composed by another ≈100 amino acids and arranged in three α-helices, corresponding to amino acids 144-154, 173-194 and 200-228 and interspersed with a antiparallel β-pleated sheet, formed by β-strands at residues 128-131, and 161-164. A single disulfide bond is found between cystein residue 179 and 214. Depending on the species, the flexible N-terminus contains at least four glycine-rich octapeptide repeats which display a particular affinity for Cu^{2+}, with reports of weaker binding to other divalent cations such as Zn^{2+}, Fe^{2+}, Ni^{2+}, and Mn^{2+}.[48] Notably, despite the low sequence identity between PrPC in chicken, turtle, or frog, and the mammalian proteins, the major structural features of PrPC are remarkably preserved in those non-mammalian species. In addition, both tertiary structure as well as the thermal stability of PrPC isolated from the brain of healthy calf were indistinguishable from bacterially expressed, recombinant

[49] Soto C. and Satani N. (2011) Trends in Molecular Medicine *17* (1), 14-24.

bovine PrPC.[50] It is unlikely that the NH_2 terminus is randomly coiled *in vivo*, and it is possible that the flexible tail of PrPC acquires a defined structure when PrPC is present within membrane rafts.[46]

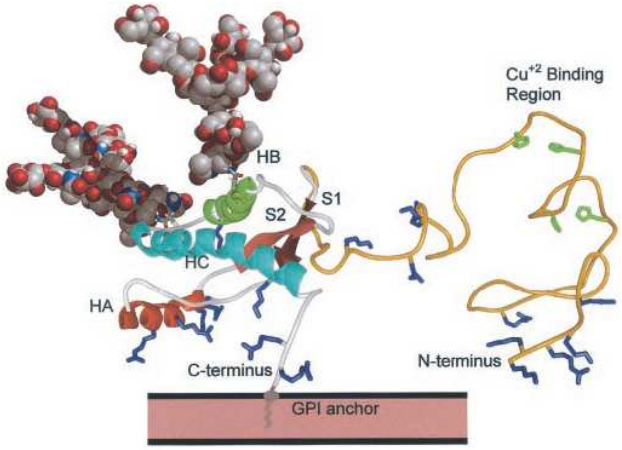


Figure 11. Model of the soluble prion protein. Residues 125–231 are from the NMR structure modeled with carbohydrates (in space-filling representation), the N-terminal region (residues 23–124) as random coil (light orange), and the GPI-membrane anchor (gray). Secondary structure is displayed as ribbons: helix A (HA; residues 144–156; red), helix B (HB; residues 177–193; green), helix C (HC; residues 200–223; cyan), β-strands S1 (residues 128–131; red), and S2 (residues 160–164; dark red). Histidine sidechains involved in copper binding are shown in green. Lysine and arginine residues are shown in stick representation (blue).

The highly conserved PrP globular domain also contains two potential sites for N-glycosilation, in fact full-length PrPC is found in non-, mono-, or diglycosylated forms, corresponding to the variable occupancy of residues Asn-181 and Asn-197 in human PrPC. A rather large variety of *N*-glycans were found attached to both full-length and truncated PrPC, which may be differentially distributed in various areas of the central nervous system. The nearly identical tridimensional structures of both recombinant PrPC and the glycosylated PrPC isolated from calf brain indicate that neither the attached *N*-glycans nor most of the GPI anchor impart major structural features by themselves. However, molecular dynamic simulations suggest that some

[50] Linden R. *et al.* (2008) Physiol. Rev. *88*, 673–728.

attached *N*-glycans may modulate PrPC stability, although experimental evidence is still lacking. Nevertheless, protein glycosylation seems to affect other protein properties, such as intracellular traffic and ligand binding. Indeed, glycosylation affects the recognition of various species of PrPC by monoclonal antibodies in both the brain and in other cells, and different responses to certain monoclonal antibodies were described for cells bearing distinctly glycosylated forms of PrPC. In turn, experimental data suggest that membrane attachment through the GPI anchor, as well as other PrPC-membrane interactions, may modulate the structure of the protein. An early study showed that PrPC tends to interact with lipid vesicles containing acidic lipids, depending on pH. Interestingly, anchorless PrPC undergoes an unusual amyloid conversion into PrPSc and aggravates disease induced in mice expressing GPI-anchored PrPC, but does not mediate clinical symptoms in infected mice lacking GPI-anchored PrPC.

Analysis of the effect of pH on the structure of the prion protein may also be relevant to physiological functions, as it was observed that acidic pH imparts some degree of structure to the N-terminal flexible domain, as well as it leads in changes in antibody binding.[50]

Although PrPC and PrPSc appear to share the same primary structure, they differ profoundly in biochemical and biophysical properties. Cellular PrPC is monomeric, proteinase-sensitive, and soluble in nonionic detergents, whereas PrPSc is insoluble and aggregate in nature, showing partial resistance to proteinase K digestion. In contrast to PrPC, the β-sheet content of PrPSc comprises ≈40% of the protein, whereas α-helices comprise ≈30% of the protein as measured by Fourier-transform infrared and CD spectroscopy. This suggests that the differences in biophysical properties between these two isoforms result from their distinct conformations. The concept that a single protein can exist in multiple

thermodynamically stable conformations would appear to challenge the Anfinsen principle that the three-dimensional structure of a protein is encoded entirely by its amino acid sequence. However, the PrPC conformational conversion into PrPSc does not necessarily violate this rule, since the process appears to be ultimately associated with polymerization of PrP, which would be expected to modulate the lowest-free energy conformation of monomeric subunits. Due to evident experimental difficulties in structural studies on large protein aggregates, such as amyloid fibrils, no high-resolution structure is available for PrPSc, although interesting models have been conjectured on the basis of electron crystallography studies.

To systematically evaluate the relationship between infectivity, converting activity, and the size of various PrPSc-containing aggregates, Caughey and coworkers[51] partially disaggregated PrPSc. The resulting species were fractionated by size and analyzed by light scattering and non-denaturing gel electrophoresis. Intracerebral inoculation of the different fractions into hamsters revealed that, with respect to PrP content, infectivity peaked markedly with 17–27 nm (300–600 kDa) particles. These results suggest that non-fibrillar particles, with masses equivalent to 14–28 PrP molecules, are the most efficient initiators of transmissible spongiform encephalopathies. As with other diseases characterized by protein aggregation, such as Alzheimer's disease and other amyloidoses, the formation of large amyloid fibrils might be a protective process that sequesters the more dangerous subfibrillar oligomers of the amyloidogenic peptide or protein into relatively innocuous deposits.[46]

The precise mechanism by which PrPSc is involved in the disease pathogenesis is largely unknown. It was observed that PrPSc accumulates in

[51] Silveira J.R. *et al*. (2005) Nature *437*, 257–261.

different brain regions as distinct types of deposits depending on the animal species and strains of the agent.[49]

According to the protein only hypothesis, there have been proposed two models to explain the conformational conversion of PrPC into PrPSc. The "template-directed refolding" hypothesis predicates an instructionist role for PrPSc on PrPC. Alternatively, the "seeded nucleation hypothesis" suggests that PrPSc exists in equilibrium with PrPC. In a non-disease state, such an equilibrium would be heavily shifted toward the PrPC conformation, such that only minute amounts of PrPSc would coexist with PrPC. If this were the case, PrPSc could not possibly represent the infectious agent, since it would be ubiquitous. According to this model, the infectious agent would consist of a highly ordered aggregate of PrPSc molecules. The aggregated state would be an intrinsic property of infectivity. Monomeric PrPSc would be harmless, but it might be prone to incorporation into nascent PrPSc aggregates. For yeast prions it was experimentally verified that an important part of their replication process is the propagation of prions through the fragmentation of existing fibrils and there are some indications that similar processes govern the growth of mammalian prions, as well as non-prion-related amyloid fibrils.

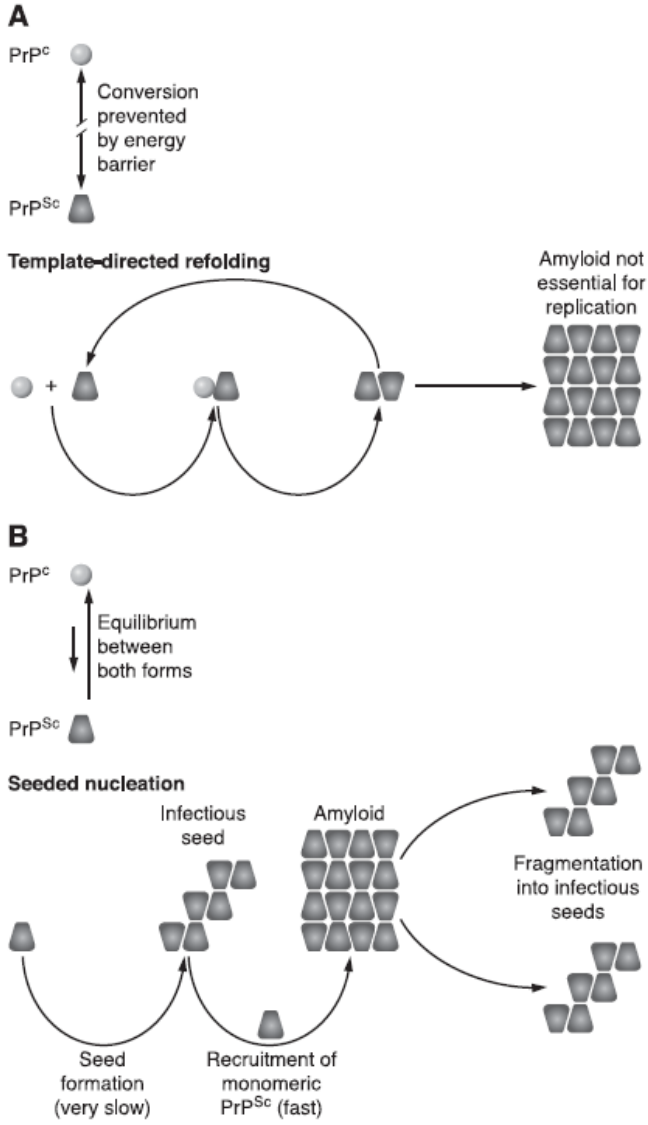


Figure 12. Models for the conformational conversion of PrPC into PrPSc. **A**: the "refolding" or template assistance model postulates an interaction between exogenously introduced PrPSc and endogenous PrPC, which is induced to transform itself into further PrPSc. A high energy barrier may prevent spontaneous conversion of PrPC into PrPSc. **B**: the "seeding" or nucleation-polymerization model proposes that PrPC and PrPSc are in a reversible thermodynamic equilibrium. Only if several monomeric PrPSc molecules are mounted into a highly ordered seed, can further monomeric PrPSc be recruited and eventually aggregate to amyloid. Fragmentation of PrPSc aggregates increases the number of nuclei, which can recruit further PrPSc and thus results in apparent replication of the agent.

In living systems, the rate constants for prion replication are clearly influenced not only by the intrinsic strength of PrPSc aggregates, but also by other cellular components: in fact, molecular chaperones in yeast, for instance, have been identified as important players in the rate of prion replication.[46]

It is still unknown how much of the whole PrPSc molecule is involved in the fibril growth. It is shown that the N-terminal part of PrP, specifically residues 112–141, can go through conformational changes involving β-strand formation, which subsequently triggers fibril growth, and solid state NMR studies showed that residues 112–141 are part of the highly ordered core of human PrP(23–144). It was previously shown that peptides based on the 89–143 region of the human PrP protein can form fibrils rich in β-sheet structure which are biologically active in transgenic mice. Within this region the human PrP(106–126) peptide is the smallest known region of PrP that forms fibrils that are toxic and resemble the physiological properties of PrPSc.[52] Indeed, PrP(106-126) polymers showed a secondary structure composed largely of β-sheet, a hydrodynamic diameter of 100 nm corresponding to the aggregation of ≈6000 molecules, and partial resistance to proteinase K and Pronase digestion. This peptide also shows neurotoxic activity toward primary cultures of rat hippocampal neurons, suggesting that the PrP region comprising residues 106-126 might give a major contributor to the structural and pathogenic properties of PrPSc and PrP amyloid.[53] This segment corresponds to an unstructured region just outside of the globular C-terminal domain of PrPC and it has been shown to adopt different secondary structures under different solution conditions,

[52] Boshuizen R.S. *et al.* (2009) The Journal Of Biological Chemistry *284* (19), 12809–12820.
[53] De Gioia L. *et al.* (1994) Journal Of Biological Chemistry *269* (11), 7859-78862.

and is thus considered as a relevant model for investigating the mechanism of fibril formation and PrPSc-mediated cell death.[54]

Despite the considerable uncertainty about the mechanism of pathogenesis in prion diseases, there is a considerable demand for both diagnostic techniques and potential therapies.[55] In the past years numerous protocols have been designed for post-mortem detection of the disease-associated prion protein or prion plaques in neural tissue[56] and there have been efforts to develop tools for ante-mortem diagnosis.

In some cases, there is no sure proof that a neurodegenerative condition is prion disease, such as CJD or BSE (Bovin Spongiform Encephalopathy), without analysis of brain tissue for the presence of spongiform changes and PrPSc. Some evidences of neurological changes or the presence of CSF markers, such as brain protein 14-3-3, can be helpful but are not specific. Tonsil biopsies in the case of variant CJD or analyses of other lymphoid tissues in the case of scrapie provide good evidence for the diagnosis of prion diseases. However, diagnosis usually can occur only when the disease has progressed to a point where the quality of life has deteriorated to such an extent that prevention of further neuronal death would have little benefit. The prion diseases all have long incubation periods, making diagnosis before clinical onset difficult. Nevertheless, there are currently considerable endeavors aimed at developing an effective test for the preclinical diagnosis of prion disease.[55] In particular blood brain barrier (BBB) permeable amyloid stains could facilitate the diagnosis of prion diseases by labeling plaques *in vivo*, using different imaging techniques.[56] Even if such a possibility is eventuated, without an effective treatment for

[54] Kuwata K. *et al.* (2003) PNAS *100* (25), 14790–14795.
[55] Webb S. *et al.* (2007) Journal Of Virology *81* (19), 10729–10741.
[56] Hoefert V.B. *et al.*(2004) Neuroscience Letters *371*, 176–180.

these diseases, such a diagnosis remains a sure death sentence. Treatment investigations target mostly the accumulation of PrPSc in the brain.[55]

A variety of compounds have been tested on cellular and animal models of prion disease. A few of these compounds, including sulphated polyanions, amphotericin B, Congo red, iododoxorubicin, tetrapyrroles, branched polyamines and modifed PrP peptides, effectively reduce the accumulation of PrPSc in scrapie-infected neuroblastoma cells and/or delay the appearance of clinical symptoms and prolong the survival time of rodents experimentally infected with scrapie; moreover it was shown that the antibiotic tetracycline binds to amyloid fibrils generated by synthetic peptides corresponding to residues 106-126 and 82-146 of human PrP, hinders assembly of these peptides into amyloid fibrils, reverts the protease resistance of PrP peptide aggregates and PrPSc extracted from brain tissue of patients with Creutzfeldt-Jakob disease and prevents neuronal death and astrocyte proliferation induced by PrP peptides *in vitro*. NMR spectroscopy revealed several through-space interactions between aromatic protons of tetracycline and side-chain protons of Ala117-119, Val121-122 and Leu125 of PrP(106-126). It can be noticed that some of these molecules (i.e. Congo red, iododoxorubicin, tetrapyrroles and tetracyiclines) share common chemicophysical properties as they all contain an extended hydrophobic core formed by aromatic moieties with a number of hydrophilic substituents conferring an amphiphilic character. Since the conversion of PrPC into PrPSc results in an increase of hydrophobicity due to solvent exposure of lipophilic residues and in the ability to form insoluble aggregates and amyloid fibrils, it is conceivable that the anti-prion activity of the compounds is related to their hydrophobic character allowing for a

strong interaction with PrPSc.[57] Despite tetracycline is considered a prototype of compounds with the potential of inactivating the pathogenic forms of PrP, impeding PrPC to PrPSc conversion, and hindering the development and progression of prion disease,[57] it has some chemical drawbacks as it lacks solubility, chemical stability and shows other pharmacological activities not directly related to amyloidosis. Trying to overcome these limitations, in our laboratory a panel of soluble, stable glycofused aromatic tricyclic compounds have been recently synthesized.[58] After having demonstrated that these molecules were able to bind Aβ-peptides,[59] they have been tested, along with other novel derivatives, against PrP(106-126) peptide and the results are reported in this thesis.

[57] Tagliavini *et al.* (2000) J. Mol. Biol. *300*, 1309-1322.
[58] Airoldi C. *et al.* (2011) Chem. Commun. *47*, 10266–10268.
[59] Airoldi C. *et al.* (2011) Org. Biomol. Chem. *9*, 463–472.

5.3 A Brief Introduction On The Field Of Biomaterials

Natural and synthetic biomaterials serve as fundamental research and therapeutic tools to investigate and facilitate the repair of damaged or dysfunctional tissues, both in cell-based and acellular therapies.[60,61,62] Many experimental and clinical applications of tissue engineering rely on the ability of stem or progenitor cells to self-organize into functional tissue.[63,64] The intrinsic genetic potential of these progenitor cells to regenerate the desired tissue acts in concert with the molecular cues present in the extracellular microenvironment to navigate the multiple differentiation pathways and produce the appropriate cell types. These progenitor cells respond to the molecular cues, which are biochemical and biophysical signals transmitted via cell surface receptors and integrated via a complex array of intracellular signaling pathways. An important part of tissue development and function relies on the dynamic dialog that exists between cells and their microenvironment. Tissue dynamics, that is, its formation, function and regeneration after damage, as well as its function in pathology, is the result of an intricate temporal and spatial coordination of numerous individual cell fate processes, each of which is induced by a myriad of signals originating from the extracellular microenvironment. (Fig. **13**)

[60] Hubbell, J. A. (2003) Curr. Opin. Biotechnol. *14*, 551.
[61] Langer, R. *et al.* (2004) Nature *428*, 487.
[62] Rahaman M. N. *et al.* (2005) Biotechnol. Bioeng. *91*, 261.
[63] Howe A. *et al.* (1998) Curr. Opin. Cell Biol. *10*, 220.
[64] Langer R. *et al.* (1993) Tissue engineering. Science *260*, 920.

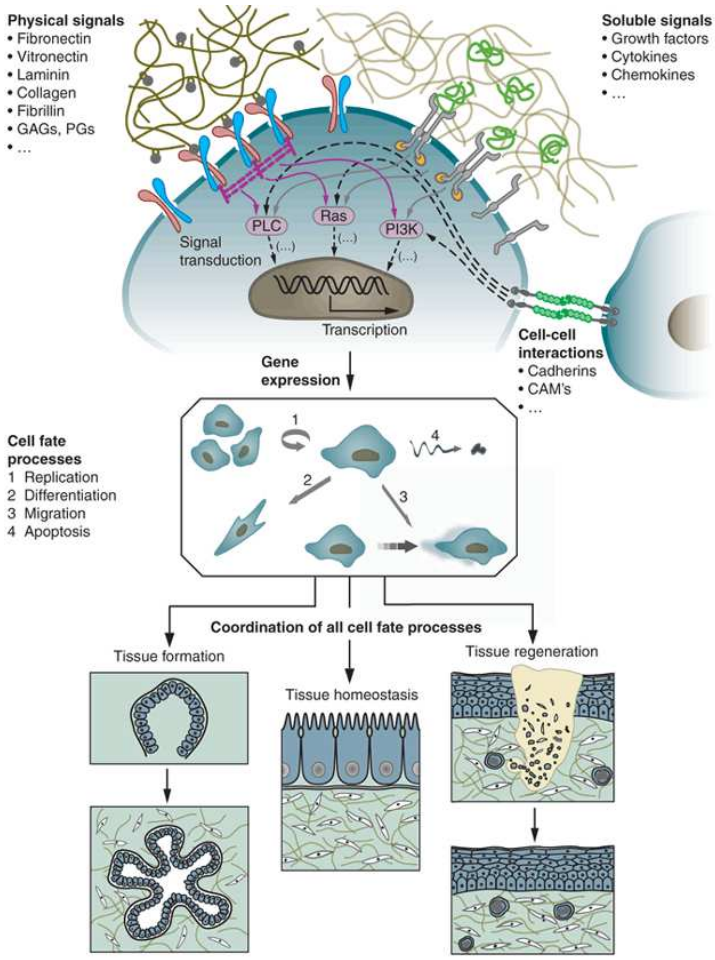


Figure 13. The extracellular microenvironment is a hydrated protein- and proteoglycan-based gel network comprising soluble and physically bound signals as well as signals arising from cell-cell interactions. Specific binding of these signaling cues with cell-surface receptors induces complex intracellular signaling cascades that converge to regulate gene expression, establish cell phenotype and direct tissue formation, homeostasis and regeneration. PLC, phospholipase C; GAGs, glycosaminoglycans; PGs, proteoglycans; CAMs, cell adhesion molecules.

In brief, a highly dynamic and complex array of biophysical and biochemical signals, transmitted from the outside of a cell by various cell surface receptors and integrated by intracellular signaling pathways, converge to regulate gene expression and ultimately establish cell phenotype. The

extracellular microenvironment, which surrounds cells and comprise the molecular cues that defines the molecular signals, is a highly hydrated network hosting three main effectors: soluble macromolecules (e.g., growth factors, chemokines, cytokines),[65] insoluble hydrated macromolecules (e.g., fibrillar proteins such as collagens, noncollagenous glycoproteins such as elastin, laminin or fibronectin, and hydrophilic proteoglycans with large glycosaminoglycan side chains) called physical signals[66,67] and proteins presented on the surface of adjacent cells (e.g., integrins, cadherins).[68] Thus, the ultimate decision of a cell to differentiate, proliferate, migrate, apoptose or perform other specific functions is a coordinated response to the molecular interactions with these ECM effectors. Environments created by isolating ECM with entrapped growth factors have been valuable tools to examine the interplay between these various molecular cues and the cellular responses they influence. Additional cues come as well as signals from surrounding cells locally degrading under the influence of proteases at the surface of the migrating cells,[69,70,71] hence promoting specific cellular processes leading to tissue formation or regeneration. The generation of smart biomaterials for tissue engineering requires mimicking natural ECMs that regulate complex morphogenetic processes in tissue formation and regeneration. Their functionality should be adjustable to a particular biological environment to obtain cell- and tissue-specificity. Ideally, one would create them from an array of biocompatible scaffolds decorated with an array of ligands inducing cell

[65] Lieberman J.R. *et al.* (2002) J. Bone Joint Surg. Am. *84*, 1032.
[66] Kogan G. *et al.*(2007) Biotechnol. Lett. *29*, 17.
[67] Morra M. (2005) Biomacromolecules *6*, 1205.
[68] Lutholf M.P. *et al.* (2005) Nat. Biotechnol. *23*, 47.
[69] Werb Z. (1997) Cell *91*, 439.
[70] Streuli C. (1999) Curr. Opin. Cell Biol. *11*, 634.
[71] Sternlicht M.D. *et al.* Annu. Rev. Cell Dev. Biol. *17*, 463.

adhesion and/or proliferation and/or differentiation. The precursor building blocks could be cross linked into solid networks by several mean. Till now, the ideal biomaterial candidate, with all these requirements, has not been projected, due to the fact that all these molecular interactions have not been definitively clarified jet. Only very few attempts to prepare biomaterial mimicking the real extracellular matrix have been proposed.[72] While artificial biomaterial matrices used in culture and tissue engineering have been developed to be responsive to physical[73,74,75]and even biochemical *stimuli*,[76,77] they do not widely address the concept to responsiveness to cellular *stimuli*.[78]

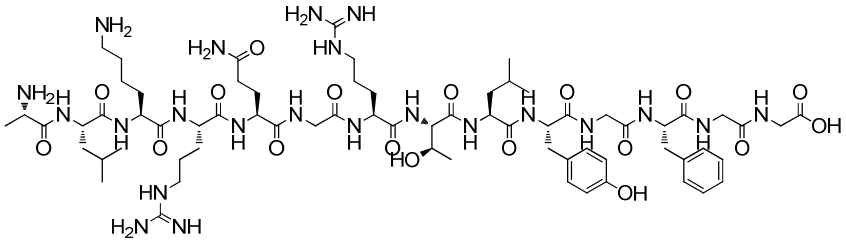
In particular, it is important to study the structural details of the topology of the different molecular interactions that occur between adhesion proteins and their small molecule counterparts (ligand). A deeper understanding of these interactions allows to project smart biomaterials properly tuned to favour the regeneration of a specific tissue. Once the nature and topology of the interaction is understood, new ligands with better, different or unique properties can be generated. Moreover it is necessary to verify that after immobilization to the material surface the ligand retains the proper conformation and therefore the ability to interact with the specific binding partners.

For tissue engineering applications it was reported that the immobilization of growth factors or peptides on tissue engineered substrates is

[72] Anderson E.H. *et al.* (2004) Macromol. Biosci. *4*, 766.
[73] Petka W.A. *et al.* (1998) Science *281*, 389.
[74] Chen G. *et al.* (1995) Nature *373*, 49.
[75] Yoshida R. *et al.* (1995) Nature *374*, 240.
[76] Miyata T. et al. (2002) Adv. Drug Delivery Rev. *54*,79.
[77] Salvaya D.M. *et al.*(2006) Mol. BioSyst. *2*, 36.
[78] Lutholf M.P. *et al.* (2003) Adv. Mater. *15*, 888.

advantageous for promotion of viability, proliferation and differentiation.[79] Among the different cues responsible for cell growth and/or differentiation, Osteogenic Growth Peptide (OGP) has recently attracted considerable clinical interest as a bone anabolic agent and hematopoietic stimulator.[80,81] Moreover osteogenic peptides have potential to be therapeutic for the treatment of osteoporosis, fracture repair and repair of loosened bone implants.[82]

OGP was discovered in the early 1990s, when it was initially isolated from regenerating bone marrow as an osteogenic cell growth promoting activity.[83] It is a highly conserved, naturally occurring tetradecapeptide, whose sequence is identical to the C-terminal amino acid sequence (89-102) of the histone H4.[84,85]



Ala-Leu-Lys-Arg-Gln-Gly-Arg-Thr-Leu-Tyr-Gly-Phe-Gly-Gly

Figure 14. The OGP sequence.

The mature bioactive form of endogenous OGP is a proteolytic cleavage product of PreOGP translated from H4 mRNA via alternative translational

[79] Moore N.M. *et al.* (2010) Biomaterials31, 1604-1611.
[80] Chen Z. *et al.* (2011) Mol Cell Biochem *358*, 221–227.
[81] Zhu H. *et al.* (2008) Veterinarni Medicina *53* (9), 501-509.
[82] Gordon E. *et al.* (2004) Curr. Med. Chem. *11*, 2599.
[83] Bab I. *et al.* (1992) Embo Journal *11*, 1867-1873.
[84] Spreafico A. *et al.* (2006) Journal of cellular biochemistry *98*, 1007-1020.
[85] Bab I. and Chorev M. (2002) Biopolymers (peptide science) *66*, 33-48.

initiation at a downstream initiation codon. OGP in high abundance occurs physiologically in human and rodent serum and in serum-free medium of osteoblastic and fibroblastic cells. The OGP small size and linearity that make OGP highly susceptible to proteolysis, and the low dose that elicit *in vivo* responses, suggested the presence of a circulating OGP-binding protein (OGPBP),[85] that maintains large reservoirs of inactive OGP protected from proteases and clearance. The OGPBP provides a mechanism controlling the peptide availability to its target cells and can be associated with the regulation of free-ligand levels and/or the ligand transport and presentation to its specific tissue or putative receptors.[86,87] The major OGPBP in serum was find to be α2-macroglobulin (α2M). α2M is a 720kDa homotetramer present in high abundance in the serum. It is a multifunctional protein that serves as a broad spectrum proteinase inhibitor and as a binding protein for several regulatory peptides. The primary circulating form of α2M is the native form; it is activated by covalently entrapping proteinases, a process that involves a conformational change that makes the molecule more compact and hence confers fast mobility when α2M is subjected to non-denaturing gel electrophoresis. The irreversible conformational change renders the α2M recognizable by its receptor, the α2M receptor/low-density lipoprotein receptor-related protein (α2M-R/LRP). This receptor is expressed in different cell types, including fibroblasts, macrophages, adipocytes and hepatocytes and presents the polypeptide regulators to their corresponding receptors, or for degradation, following internalization of the complex.[88] It was demonstrated that both the native and the activated forms of α2M bind OGP noncovalently; moreover the native α2M seems to significantly enhance the OGP mitogenic activity in osteoblastic

[86] Greenberg Z. *et al.* (1997) Journal of cellular biochemistry *63*, 359-367.
[87] Greenberg Z. *et al.* (1993) Biochmica and Biophysica Acta *1178*, 273-280.
[88] Gavish H. *et al.* (1997) Biochemistry *36*, 14883-14888.

MC3T3 E1 cells, whereas this activity seems to be inhibited by activated α2M.[88] Dissociation of OGP from its complex with α2M, exposes OGP to proteolytic cleavage, thus probably generating the C-terminal pentapeptide OGP(10–14), which was shown to activate at least an intracellular Gi-protein-MAP kinase signaling pathway.[89] OGP(10-14) is indeed the minimal amino acid sequence that retains the full OGP-like activity, but seems to lack the ability to form OGP-OGPBP complexes.[90] Tyr-10, Phe-11, Gly-13 and Gly-14 are essential for the OGP(10-14) activity. Moreover, through the design of truncated and conformationally constrained analogues, it was suggested that the spatial presentation of the Tyr and Phe side chains has a major role in the productive interaction with the putative receptor.[91]

For most tissue enigineering applications, immobilized OGP is more practical than soluble OGP for maintaining tight spatial proximity with the cells at the wound site for extended periods of time. OGP has been already incorporated into peptide nanofiber scaffolds or Poly Lactic-co-Glycolic Acid (PLGA) gels and it was shown that it increased cell density compared to control substrates. OGP and OGP(10-14) have been also coupled to oxidized self assembled monolayer gradients by a polyethylene oxide linker using click-chemistry[92] and it was observed that similar to soluble OGP, immobilized OGP increased cell proliferation of osteoblast lineage cells, but the increases were independent of immobilized OGP concentration, probably because immobilized OGP may only be active when the OGP(10-14) region is cleaved from the substrate surface. This was supported by the observation that immobilized OGP(10-14) had a lower impact on cell proliferation compared to the full peptide, suggesting that the active region

[89] Gabet Y. *et al.* (2004) Bone *35*, 65-73.
[90] Chen Y-C. *et al.* (2000) J.Peptide research *56* 147-156.
[91] Chen Y-C. *et al.* (2002) J.Med.Chem. *45*, 1624-1632.
[92] Hartmuth C. *et al.* (2001) Angew. Chem. Int. Ed. *40*, 2004-2021.

needs to be cleaved from the substrate to enhance cell proliferation. These initial results are therefore promising towards the development of novel biomaterials in which OGP is linked as a proliferative agent.[79]

Other important actors, playing a pivotal role in promoting cell adhesion and differentiation and, thus, tissue regeneration, are carbohydrates and lectins. Carbohydrates are by far the most abundant organic molecules found in nature, and nearly all organisms synthesize and metabolize carbohydrates.[93] They are polyhydroxy aldehydes or ketones, or substances that yield such compounds on hydrolysis. There are three major size classes of carbohydrates: monosaccharides, oligosaccharides, and polysaccharides (the word "saccharide" is derived from the Greek *sakcharon,* meaning "sugar"). Monosaccharides, or simple sugars, consist of a single polyhydroxy aldehyde or ketone unit; oligosaccharides consist of short chains of monosaccharide units, or residues, joined by characteristic linkages called glycosidic bonds, whereas the polysaccharides are sugar polymers containing more than 20 or so monosaccharide units, and some have hundreds or thousands of units. In cells, most oligosaccharides, consisting of three or more units, do not occur as free entities but are joined to nonsugar molecules (lipids or proteins) in glycoconjugates.[94] The complex heterogeneity of carbohydrates in living systems is a direct result of several carbohydrate characteristics: the ability of different types and numbers of sugar residues to form glycosidic bonds with one another, the structural characteristics of these molecules, the type of anomeric linkage, the position and the absence or presence of branching. The structural variability and complexity of cell surface glycans allows them to act as signaling molecules, recognition molecules and adhesion molecules.

[93] Ghazarian H. *et al.* (2011) acta histochemica 113 (2011) 236–247.
[94] Lehninger, Biochemistry, 4th edition, chapter 7, 238.

Indeed, cell surface glycans are involved in many physiologically important functions that include normal embryonic development, differentiation, growth, contact inhibition, cell–cell recognition, cell signaling, host–pathogen interaction during infection, host immune response, disease development, metastasis, intracellular trafficking and localization, rate of degradation and membrane rigidity. It is difficult to discuss carbohydrates without reference to lectins.[93] Lectins are defined as proteins that preferentially recognize and bind carbohydrate complexes protruding from glycolipids and glycoproteins.[95] The term lectin is derived from *lectus*, the past participle of the Latin word *legere* meaning "to choose" or "select", and has been generalized to encompass all non-immune carbohydrate-specific agglutinins regardless of blood-type specificity or source. Lectin interaction with particular carbohydrates can be as specific as the interaction between an antigen and an antibody or an enzyme and its substrate. Lectins bind not only to oligosaccharides on cells but also to free-floating glycans, including monosaccharides. Structural studies conducted on animal lectins suggest that the carbohydrate-binding activity of most lectins depends on a limited number of aminoacid residues designated as the carbohydrate recognition domain (CRD). The CRD typically recognizes the terminal non-reducing carbohydrate residues of cell membrane glycoproteins and glycolipids. Lectin CRDs also may discriminate between anomeric isomers, as a function of their specificities.[96,97]

Commonly, lectins have been classified according to sequence similarity and tertiary structure, but it is more convenient to classify carbohydrate-binding proteins according to the topological features of the binding site to

[95] Ambrosi M. *et al.* (2005) Org. Biomol. Chem. *3*, 1593-1608.
[96] Mody R. *et al.* (1995) J. Pharmacol. Toxicol. Methods *33*, 1–10.
[97] Ernst B. and Magnani J.L.(2009) Nature Reviews-Drug discovery *8*, 661-677.

detect the possible common principles underlying carbohydrate recognition. Carbohydrate-binding proteins can be divided in two major groups. Group I is constituted by the proteins that have a buried binding site and engulf the ligand fully upon the binding like the bacterial periplasmatic transport proteins and enzymes. Group II is formed by proteins that have a shallow binding site, mostly in the form of a depression on the protein surface. Three classical lectin families belong to this group: legume lectins, C-type lectins and galectins. C-type lectins (CTLs) are the most abundant of all animal lectins, and the CTL superfamily is grouped into three families: selectines, collectins and endocytic lectins. A majority of CTLs are large and asymmetric, have one or more CRDs and exist as Ca^{2+}-dependent proteins found in secreted or bound forms. Instead, the galectins in the CTL superfamily are generally small, non-glycosylated, soluble and exist as Ca^{2+}-independent intracellular proteins.[98]

One of the most important characteristic of lectins is the ability to discriminate between Galactose (Gal) and Glucose (Glc) or Mannose (Man) residues; as a matter of fact, being lectins highly selective, they can distinguish between carbohydrates that differ only in the stereochemistry at one carbon atom (epimers). The physicochemical nature of sugar-protein interaction has been a matter of debate for years. It has been established that the single molecule sugar-protein interactions are weak in nature and it has been suggested that multivalency is a key feature for the molecular recognition process to take place. Because of the amphipatic character of the oligosaccharide, different kinds of forces may be involved in its recognition by receptors: the presence of the hydroxyl groups obviously makes possible their involvement in intermolecular hydrogen bonds to side chains of polar amino acids and it has also been proposed that water

[98] Elgavish S. and Shaanan B. (1997) Trends Biochem Sci. *22* (12), 462-467.

provides the driving force for the intermolecular interaction. Nevertheless, not only polar forces are involved in carbohydrate recognition. NMR and X-ray diffraction data have shown that, depending on the stereochemistry of the saccharide, many rather nonpolar C-H groups indeed lead to the formation of patches that interact with the aromatic residues of protein side chains.[99,100,101]

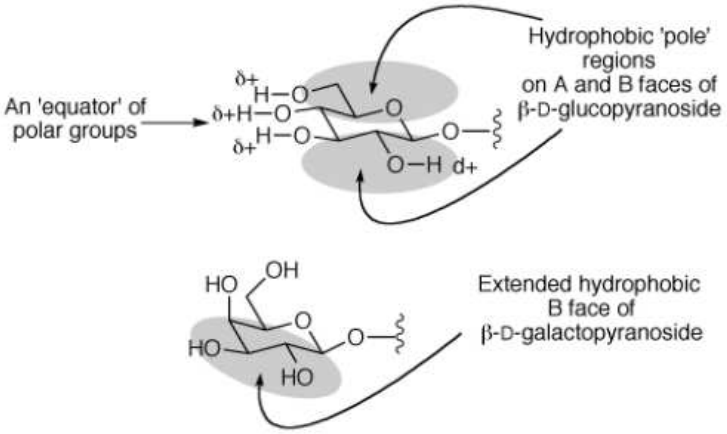


Figure 15. Hydrophobic and polar patches on carbohydrates.

Hence, both non polar interactions and hydrogen-bonding play important roles in the binding process. In the last few years, the presence of aromatic rings in the binding sites of lectins has been highlighted as essential for recognition of neutral sugars, like Gal/GalNAc and Glc/GlcNAc families. Both the structure and the conformation of the carbohydrate, as well as the nature and orientation of the aromatic rings, are important for the molecular recognition process. This so-called "stacking interaction" between carbohydrates and aromatic amino acid side chains is also referred as "CH-π interaction" and, in order to better understand this kind of

[99] Fernandez-Alonso M. C. *et al.* (2005) J. Am. Chem. Soc. 127, 7379-7886.
[100] Vandenbussche S. *et al.* (2008) Chem. Eur. J. *14*, 7570 – 7578.
[101] Asensio J.L. *et al.* Acc. Chem. Res., Article ASAP DOI: 10.1021/ar300024d

interactions, some groups have studied both experimentally and theoretically simple models of interactions between different D-monosaccharides and several aromatic entities. These studies demonstrate that π-electron-rich aromatic rings seem to be required in order to establish stabilizing CH-π interactions with carbohydrates, and all suggest that three C-H vectors pointing into the same spatial region need to exist to allow that the interaction takes place.[99,100,101]

Results and Discussion

Objective of the thesis

NMR techniques allow to obtain fine structural information useful for the comprehension of biological processes. This thesis collects the results obtained through the application of high resolution NMR in solution and HR-MAS techniques to the characterization of different molecular recognition processes, occurring between biological macromolecules and binding partners, such as other macromolecules or small ligands. This knowledge was and will be exploited towards the development of new therapeutic and diagnostic strategies.

Chem. Eur. J. 2010, 16, 1897 – 1902

Targeting Bacterial Membranes: NMR Spectroscopy Characterization of Substrate Recognition and Binding Requirements of D-Arabinose-5-Phosphate Isomerase

Cristina Airoldi, Silvia Sommaruga, Silvia Merlo, Paola Sperandeo, Laura Cipolla, Alessandra Polissi, and Francesco Nicotra

Abstract: Lipopolysaccharide (LPS) is an essential component of the outer membrane of Gram-negative bacteria and consists of three elements: lipid A, the core oligosaccharide, and the *O*-antigen. The inner-core region is highly conserved and contains at least one residue of 3-deoxy- D-manno-octulosonate (Kdo). Arabinose-5-phosphate isomerase (API) is an aldo–keto isomerase catalyzing the reversible isomerization of D-ribulose-5-phosphate (Ru5P) to D-arabinose-5-phosphate (A5P), the first step of Kdo biosynthesis. By exploiting saturation transfer difference (STD) NMR spectroscopy, the structural requirements necessary for API substrate recognition and binding were identified, with the aim of designing new API inhibitors. In addition, simple experimental conditions for the STD experiments to perform a fast, robust, and efficient screening of small libraries of potential API inhibitors, allowing the identification of new potential leads, were set up. Due to the essential role of API enzymes in LPS biosynthesis and Gram-negative bacteria survival, by exploiting these data, a new generation of potent antibacterial drugs could be developed.

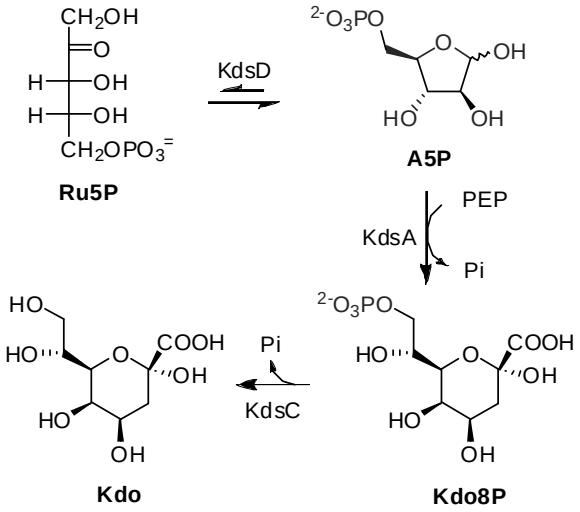
Keywords: antibiotics · enzymes · inhibitors · isomerization · NMR spectroscopy

Introduction

Arabinose-5-phosphate isomerase (API) is a key enzyme in the biosynthesis of lipopolysaccharide (LPS), which is an essential component of the outer

membrane (OM) of Gram negative bacteria.[1] Due to the presence of LPS, the OM represents an effective permeability barrier that prevents the entry of hydrophobic molecules, such as bile salts, detergents, and lipophilic antibiotics, in the cell.[2] Moreover LPS, also known as endotoxin, is one of the major virulence factors of Gram-negative bacteria and is responsible for the activation of the mammalian innate immune response.[3] LPS is a complex glycolipid that can be structurally divided into three regions: the OM-embedded lipid A; the central oligosaccharide core region; and a polysaccharide chain, known as *O*-antigen. The core provides the link between the highly conserved membrane-embedded lipid A and the structurally diverse *O*-antigen chain and can be further divided into an inner and outer core.[1, 4] Whereas the structure of the outer core is somewhat variable, the inner-core region is fairly conserved with one structural element, the 3-deoxy- D-manno-octulosonate (Kdo) residue, always present in all inner cores analyzed so far.[5] API is an aldo-keto isomerase catalyzing the first reaction of Kdo biosynthesis, namely, the conversion of D-ribulose-5-phosphate (Ru5P) into D-arabinose-5-phosphate (A5P) and represents a key target for controlling the pathway (Scheme 1).

LPS is essential for the viability of most Gram-negative bacteria, with the notable exception of Neisseria meningitides,[6] and the Kdo2-lipid A moiety represents the minimal structure indispensable for growth.[1] In fact, interruption of Kdo biosynthesis leads to the arrest of cell growth.[7] Thus, despite the fact that the requirement for Kdo may be overcome by several suppressor mutations,[8, 9] its biosynthesis represents an ideal target for antibiotic discovery.



Scheme 1. The Kdo biosynthetic pathway. API (encoded in *E. coli* by the *kdsD* gene) catalyses the first reaction of this metabolic pathway. PEP=phosphoenolpyruvate, P_i=inorganic phosphate.

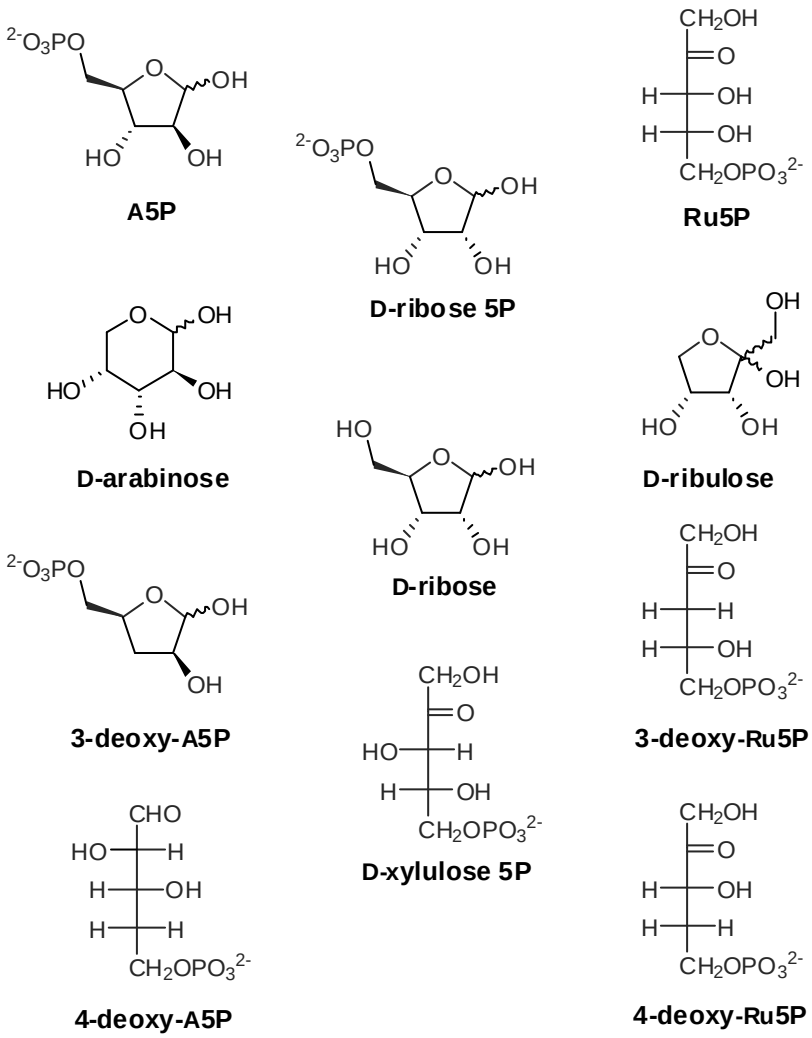
In *Escherichia coli*, API is encoded by two paralogous genes *kdsD* and *gutQ*, which may substitute for each other.[10–12] A third paralogous copy of the API enzyme, KpsF, is also present in pathogenic *E. coli* strains. Paralogous copies of API have a narrow phylogenetic distribution, with GutQ being confined to a subset of enterobacteria and KpsF present in pathogenic *E. coli* strains synthesizing group 2 capsules.[8, 10] The three E. coli APIs share greater than 40% sequence identity, as well as the same domain architecture;[8] therefore, drugs acting against API should be effective against paralogous enzymes. In fact, KdsD, GutQ, and KpsF are all tetrameric[8] and possess the sugar isomerase (SIS) domain, found in a variety of proteins that have a common role in phosphosugar isomerization, followed by a pair of cystathionine-β-synthase (CBS) domains of unknown function.[13] Although the 3D structure of KdsD has recently been predicted by homology modeling,[14] no details of the catalytic mechanism of the interconversion of Ru5P to A5P have been elucidated. On this basis,

we describe herein, the identification of the structural requirements necessary for API natural substrate recognition and binding, with the aim of designing new API inhibitors.

Results and Discussion

Saturation transfer difference (STD) NMR spectroscopy has been used extensively as a comprehensive and efficient method to investigate enzyme–ligand interactions.[15–20] We have exploited STD NMR experiments to elucidate the binding between KdsD, its natural substrates, and a number of analogues (Scheme 2), some of which have been synthesized by us and will be reported in due course.

The success of STD experiments requires that the dissociation rate of the ligand is faster than its relaxation rate (1/ T_1).[21] For an enzyme–substrate pair, this condition is satisfied for a medium–high off-rate of the enzyme–substrate complex (k_{-1}) and a low enzyme turnover number (k_{cat}). Thus, in a first step, the kinetics of the enzyme-catalyzed isomerization was monitored by ^{1}H NMR spectroscopy to find the experimental conditions (buffer composition, pH, and temperature) in which the kinetic rate was suitable for STD experiments (Figure 1S in the Supporting Information). KdsD specific activity, determined as a decrease of the Ru5P H3 integral value versus time per mg of KdsD, was a hundredfold lower at 5°C ((4.14±0.096)x10^{-4} at 37°C versus (6.58±0.377)x10^{-6} at 5°C). To prove that the KdsD secondary structure was retained after the temperature shift from 37 °C (physiological temperature) to 5°C, the circular dichroism (CD) spectra of the protein recorded at the two temperatures were compared and were identical (Figure 2S in the Supporting Information).



Scheme 2. A small library of natural KdsD substrates and analogues, the binding of which to the enzyme has been investigated by STD-NMR spectroscopy.

Based on these data, we carried out the experiments by dissolving our samples in phosphate buffer saline (PBS), pH 7.3, at 5°C. STD spectra of the natural substrates in the presence of KdsD were acquired after the mixture reached the thermodynamic equilibrium (A5P/Ru5P, 7:3).

We verified that under these experimental conditions substrate interconversion does not affect the measured STD effects. The data clearly indicated a major involvement of the protons at C-1 and C-3 in the interaction with the enzyme (Figure 1).

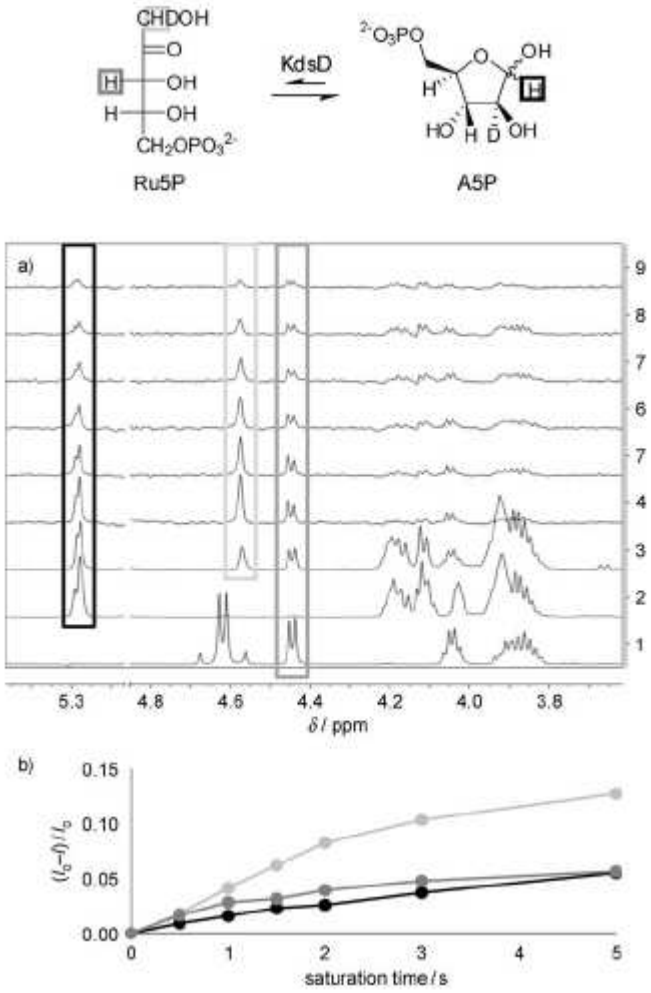


Figure 1. The reversible isomerization of Ru5P to A5P catalyzed by KdsD. When the reaction is carried out in D_2O, deuteration of Ru5P occurs at position 1 and at position 2 of A5P. As a consequence, the Ru5P H1 signal becomes a singlet. a) ^{1}H NMR spectra of Ru5P (1), A5P (2), and KdsD–substrate mixture at equilibrium (A5P/Ru5P 7:3) (3), number of scans (NS)=16; STD spectra of KdsD–substrate mixture at equilibrium, at a KdsD/A5P 1:230 molar ratio and a KdsD/Ru5P 1:100 molar ratio, recorded with different protein saturation times: 5 (4); 3 (5); 2 (6); 1.5 (7); 1 (8); 0.5 s (9), NS=768, on-resonance frequency= 7.9 ppm, off-resonance frequency=40 ppm. Spectra 3–9 were recorded on the same sample; all samples were

dissolved in phosphate-buffered saline (PBS), pH 7.3, 5°C. b) Fractional STD effects for A5P H1 (black), Ru5P H1 (light gray) and Ru5P H3 (gray) calculated by $(I_0\text{-}I)/I_0$, in which $(I_0\text{-}I)$ is the peak intensity in the STD spectrum and I_0 is the peak intensity in the off-resonance spectrum.

Although the protons in position 1 are directly involved in the isomerization mechanism, and thus, their proximity to the enzyme catalytic site was predictable, the prominent role played by the C-3 region and its stereochemistry was much more interesting. Hence, the role of position 3 was examined in more detail. STD experiments on D-xylulose-5-phosphate (Xu5P), the D-ribulose 5-phosphate epimer at position 3, revealed that Xu5P does not bind to KdsD, suggesting that inversion of configuration at position 3 inhibits recognition and binding processes (Figure 2).

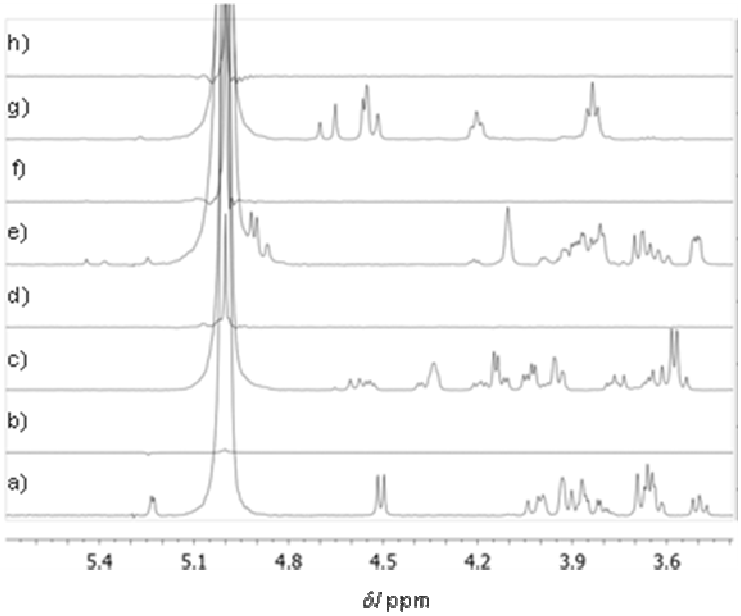


Figure 2. a) ^{1}H NMR spectrum of KdsD and D-arabinose mixture, NS= 16. b) STD spectrum of KdsD and D-arabinose mixture. c) ^{1}H NMR spectrum of KdsD and D-ribulose mixture, NS=16. d) STD spectrum of KdsD and D-ribulose mixture. e) ^{1}H NMR spectrum of KdsD and D-ribose mixture, NS=16. f) STD spectrum of KdsD and D-ribose mixture. g) ^{1}H NMR spectrum of KdsD and D-xylulose-5P mixture, NS=16. h) STD spectrum of KdsD and D-xylulose-5P mixture. KdsD/tested compound 1:100 molar ratio. For all of the STD experiments: NS=768; protein saturation time=2 s, on-resonance frequency=7.9 ppm, off-resonance frequency=40 ppm. All the samples were dissolved in PBS, pH 7.3, 5°C.

To further investigate the role of 3-OH, 3-deoxy-D-*threo*pentofuranose 5-phosphate (3-deoxy-A5P) was synthesized and the ability of KdsD to accept this compound as a substrate was tested. KdsD converts 3-deoxy-A5P to 3-deoxy-D-*glycero*-pentulose 5-phosphate (3-deoxy-Ru5P), but the catalysis proceeds tenfold slower than for the natural substrate isomerization (Table 1).

Table 1. KdsD specific activity for the different substrates at 37°C.

Substrate	Product	Aldose/ketose ratio at equilibrium	KdsD specific activity[a]
A5P	Ru5P	7:3	(2.306±0.467) $x10^{-4}$
3-deoxy-A5P	3-deoxy-Ru5P	4:6	(3.295±0.165)$x10^{-5}$
4-deoxy-A5P	4-deoxy-Ru5P	1:9	(4.187±0.491) $x10^{-4}$

[a] Specific activity was determined as an increase of H3 (Ru5P and 3-deoxy-Ru5P) or H4 (4-deoxy-Ru5P) integral value versus time per ug of KdsD. Each measurement is the mean of at least five independent determinations.

Also, STD intensities were significantly lower, as a consequence of a decrease in affinity (Figure 3). These data indicated that 3-OH deletion reduces the binding affinity, even if it does not inhibit isomerization.

The phosphate group at position 5 is also essential for the substrate interaction with the enzyme. In fact, KdsD does not bind nonphosphorylated substrate analogues, as deduced by STD data gathered on D-arabinose, D-ribose, and D-ribulose (Figure 2). On the other hand, D-ribose-5-phosphate (R5P) is recognized by the enzyme (Figure 4), but it is not converted to the corresponding ketose. The ability of R5P to compete with the natural substrates at the KdsD active site was demonstrated by competitive STD titration experiments.

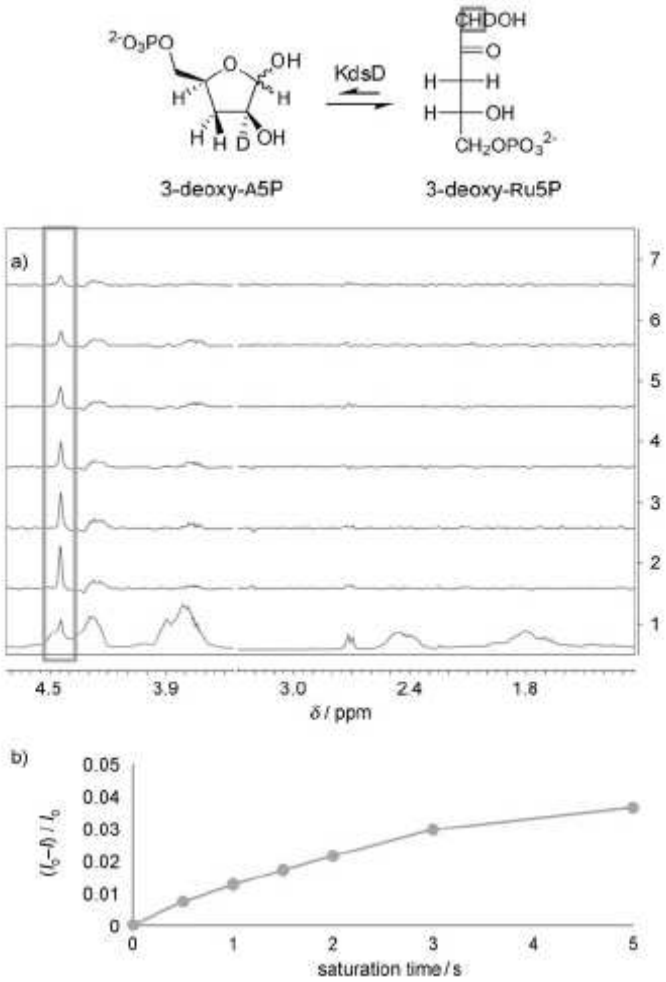


Figure 3. The reversible isomerization of 3-deoxy-Ru5P to 3-deoxy-A5P catalyzed by KdsD. When the reaction is carried out in D_2O, the deuteration of 3-deoxy-Ru5P and 3-deoxy-A5P occurs at positions 1 and 2, respectively. a) ^1H NMR spectrum of KdsD and 3-deoxy-substrate mixture at equilibrium (3-deoxy-A5P/3-deoxy-Ru5P 4:6) NS=16 (1); STD spectra of KdsD–substrate mixture at equilibrium, at a KdsD/3-deoxy-A5P 1:80 molar ratio and a KdsD/3-deoxy-Ru5P 1:120 molar ratio, recorded with different protein saturation times: 5 (2); 3 (3); 2 (4); 1.5 (5); 1 (6); 0.5 s (7), NS=2304, on-resonance frequency=7.9 ppm, off-resonance frequency=40 ppm. Spectra were recorded on the same sample; the sample was dissolved in PBS, pH 7.3, 5°C. b) Fractional STD effects for 3-deoxy- Ru5P H1 calculated by $(I_0\text{-}I)/I_0$, in which $(I_0\text{-}I)$ is the peak intensity in the STD spectrum and I_0 is the peak intensity in the off-resonance spectrum.

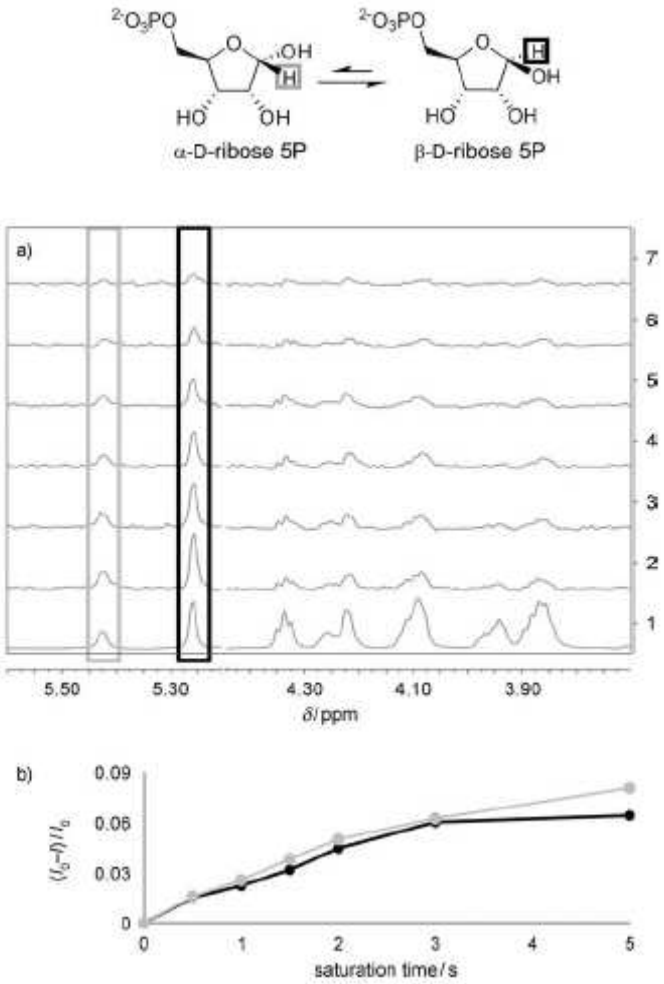


Figure 4. R5P exists in solution as a mixture of α and β anomers. a) ^{1}H NMR spectrum of KdsD and R5P mixture NS=16 (1); STD spectra of KdsD–R5P mixture, at a KdsD/R5P 1:100 molar ratio, recorded with different protein saturation times: 5 (2); 3 (3); 2 (4); 1.5 (5); 1 (6); 0.5 s (7), NS=768, on-resonance frequency=7.9 ppm, off-resonance frequency= 40 ppm. Spectra were recorded on the same sample. The sample was dissolved in PBS, pH 7.3, 5°C. b) Fractional STD effects for R5P H1 calculated by $(I_0\text{-}I)/I_0$, in which $(I_0\text{-}I)$ is the peak intensity in the STD spectrum and I_0 is the peak intensity in the off-resonance spectrum.

By increasing R5P concentration, a decrease in the substrate STD intensities was observed, particularly for Ru5P H1 (Figure 5 and Supporting Information, Figure 3S).

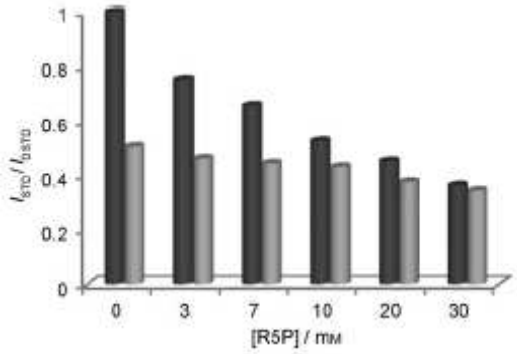


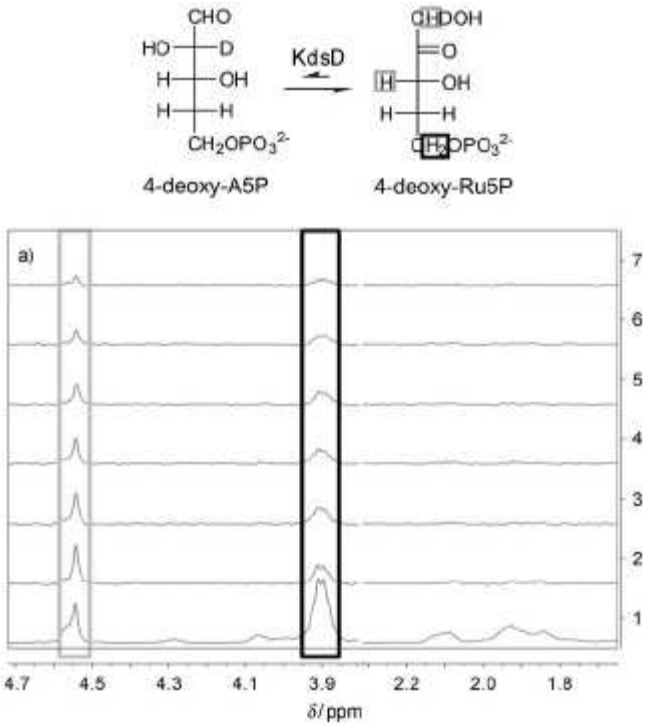
Figure 5. The decrease in the fractional STD effects of Ru5P H1 (dark gray), A5P H1 (light gray) as a function of R5P concentration. A5P, Ru5P, and KdsD concentrations were, respectively, 7 mM, 3mM, and 30 mM. Fractional STD effects were calculated by $(I_0\text{-}I)/I_0$, in which $(I_0\text{-}I)$ is the peak intensity in the STD spectrum and I_0 is the peak intensity in the off-resonance spectrum. The largest STD effect (I_{0STD}) was set to one and the relative intensities (I_{STD}) were determined.

To elucidate the importance of position 4 for substrate recognition and transformation, 4-deoxy-D-*threo*-pentafuranose-5-phosphate (4-deoxy-A5P) was synthesized as previously reported.[22] Due to the lack of a 4-OH group, 4- deoxy-A5P is unable to form the cyclic hemiacetal and retains an open-chain geometry in solution. Hence, we exploited this compound to verify if the aldose substrate was recognized preferentially in the cyclic or open-chain form. By incubating 4-deoxy-A5P in the presence of KdsD, the formation of 4-deoxy-D-*glycero*-pentulose-5-phosphate (4-deoxy-Ru5P) was observed. The isomerization rate for this compound was twofold higher than that of the natural substrate, showing a preference for the open-chain analogue (Table 1).

STD experiments confirmed the ability of KdsD to recognize and bind 4-deoxy-A5P and 4-deoxy-Ru5P (Figure 6). Therefore, the presence of a hydroxyl group at position 4 is not essential for recognition and binding of substrates or for enzymatic catalysis. It has been described that late transition metals (Ni^{2+}, Cu^{2+}, Zn^{2+}, Cd^{2+}, Hg^{2+}) inhibit KdsD, particularly those with the d^{10} electron configuration. IC_{50} for Zn^{2+} was estimated to be

between 1 and 3 mM.[11] Thus, we investigated the effect of Zn^{2+} on substrate binding to KdsD, by performing STD experiments in the presence of different $ZnCl_2$ concentrations.

According to our data, Zn^{2+} indeed inhibited substrate binding, but in a reversible manner. In fact, Zn^{2+} addition caused a substantial decrease in the substrate STD intensities. However, a few minutes after incubation with an excess of ethylenediaminetetraacetic acid (EDTA), substrate STD intensities were completely restored (Figure 7).



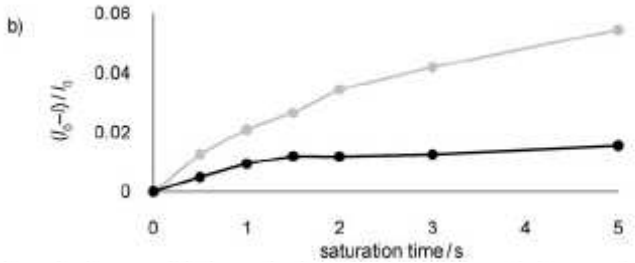


Figure 6. The reversible isomerization of 4-deoxy-A5P to 4-deoxy-Ru5P catalyzed by KdsD. When the reaction is carried out in D_2O, the deuteration of 4-deoxy-Ru5P and 4-deoxy-A5P occurs at positions 1 and 2, respectively. a) ^{1}H NMR spectrum of KdsD and 4-deoxy-substrate mixture at equilibrium (4-deoxyA5P/4-deoxy-Ru5P 1:9) NS=16 (1); STD spectra of KdsD and substrate mixture at equilibrium, at a KdsD/4 deoxy-A5P 1:17 molar ratio and a KdsD/4-deoxy-Ru5P 1:150 molar ratio, recorded with different protein saturation times 5 (2); 3 (3); 2 (4); 1.5 (5); 1 (6); 0.5 s (7), NS=768, on-resonance frequency=7.9 ppm, off-resonance frequency= 40 ppm. Spectra were recorded on the same sample; the sample was dissolved in PBS, pH 7.3, 5°C. b) Fractional STD effects for 4-deoxy-Ru5P H1 and 4-deoxy-Ru5P H3 (light gray), and for 4-deoxy-Ru5P H5 (dark gray) calculated by $(I_0\text{-}I)/I_0$, in which $(I_0\text{-}I)$ is the peak intensity in the STD spectrum and I_0 is the peak intensity in the off resonance spectrum.

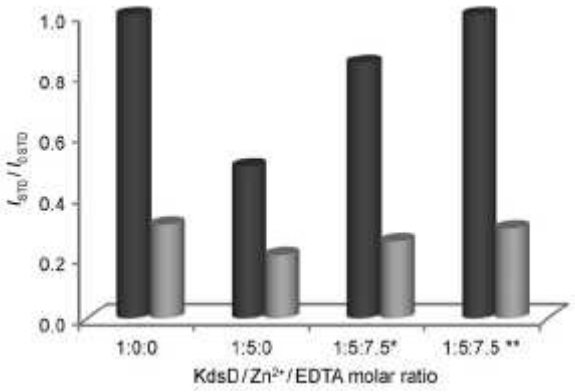


Figure 7. The decrease of the fractional STD effects of Ru5P H1 (dark gray) and A5P H1 (light gray) as a function of KdsD/Zn^{2+} molar ratio. Fractional STD effects were calculated by $(I_0\text{-}I)/I_0$, in which $(I_0\text{-}I)$ is the peak intensity in the STD spectrum and I_0 is the peak intensity in the off-resonance spectrum. The largest STD effect (I_{0STD}) was set to 1 and the relative intensities (I_{STD}) were determined. * STD experiment was performed immediately after EDTA addition; ** STD experiment was carried out 30 min after EDTA addition.

Conclusion

The importance of the molecular region encompassing positions 1 and 3, as well as the 5-phosphate group, for the binding of the natural substrates to KdsD has been demonstrated. Our data provide atomic resolution because the regions defined by substituents at positions 2 and 4 do not seem to be essential for the interaction with the enzyme. Thus, they can be chemically modified to obtain substrate analogues with potential inhibitory activity. The results obtained may be exploited in different ways: first, the simple experimental conditions for the STD experiments can be employed to perform a fast, robust, and efficient screening of small libraries of potential API inhibitors, allowing the identification of new potential leads. Second, these new structural data might be useful for the rational drug design of novel KdsD inhibitors, which, due to the essential role of API enzymes for LPS biosynthesis and Gram-negative bacteria survival, represent a new generation of potent antibacterial drugs. Third, our experimental protocol, based on the characterization of natural substrate–KdsD interaction by STD, may be applied in a general manner, since it provides an efficient method for the investigation of the substrate–enzyme binding and recognition processes, which can be extended to other enzyme–substrate pairs.

Experimental Section

Protein expression and purification: Full-length KdsD (NP_417664.1) was amplified by polymerase chain reaction (PCR) from *E. coli* K12 genomic DNA and inserted into pET23d, which allows the expression of wild-type KdsD without any tag sequence. The expression construct was initially transferred into DH5α cells for plasmid amplification and then to BL21(DE3)pLysE cells for protein expression. Wild-type KdsD was purified by Ni-chelate affinity chromatography, exploiting the specific binding of KdsD

to this resin. *E. coli* cells were grown under shaking at 37°C in LD medium containing 100 ugmL^{-1} ampicillin and 25 ugmL^{-1} chloramphenicol, until OD_{600} reached 0.6. Induction was then carried out for 2 h with 0.5 mM IPTG. Cells were harvested, washed with sodium phosphate (20 mM, pH 7.0), and resuspended in three volumes of sodium phosphate (50 mM, pH 8.0), NaCl (300 mM), and PMSF (1 mM). Cells were disrupted by sonication and centrifuged at 40000 *g* for 30 min. The supernatant was loaded at a flow rate of 0.5 mLmin^{-1} onto a Ni-NTA Agarose (Qiagen) column (bed volume 5 mL) pre-equilibrated with sodium phosphate (50 mM, pH 8.0) and NaCl (300 mM). The column was washed with 10 volumes of sodium phosphate (50 mM, pH 8.0), NaCl (300 mM), and imidazole (10 mM) and the enzyme was eluted with a 50 mL stepwise imidazole gradient (100 to 300 mM) in the same buffer. Fractions (5 mL) were collected and those displaying the highest purity (>95%, as assessed by SDS-PAGE analysis) were dialyzed in sodium phosphate (50 mM, pH 8.0) and NaCl (300 mM). Protein aliquots (600 ug each) were lyophilized and stored at-80°C.

NMR spectroscopy experiments: All of the experiments were recorded on a Varian 400 MHz instrument. The ligand resonances were assigned by using $^{1}H–^{1}H$ COSY and $^{1}H–^{13}C$ HSQC NMR spectroscopy. A basic 1D-STD sequence was used with the on-resonance frequency of 7.9 ppm and the off-resonance frequency of 40 ppm. A train of Gaussian-shaped pulses of 50 ms each was employed, with a total saturation time of the protein envelope of 5, 3, 2, 1.5, 1, or 0.5 s. The total saturation time was adjusted by the number of shaped pulses. A T1ρ filter of 2 ms was employed to eliminate the background signals from the protein. All of the samples were dissolved in deuterated PBS at pH 7.3, containing NaCl (300 mM), at 5°C. Total sample volumes were 550 uL. The on- and off-resonance spectra were acquired

simultaneously with the same number of scans. The STD spectrum was obtained by subtraction of the on-resonance spectrum from the off resonance spectrum. Subtraction was performed by phase cycling to minimize artifacts arising from magnet and temperature instabilities.[23] Reference experiments of samples containing only the free compounds tested were performed under the same experimental conditions to verify true ligand binding. The effects observed in the presence of the protein were due to true saturation transfer, as no signal was present in the STD spectra obtained in the reference experiments, except residues from HDO, indicating artifacts from the subtraction of compound signals to be negligible. API specific activity was determined as a decrease of the Ru5P H3 integral value versus time per mg of KdsD or as an increase of the H3 (Ru5P and 3-deoxy-Ru5P) or H4 (4-deoxy-Ru5P) integral value versus time per mg of KdsD. Spectra were collected over 673 s at intervals of 33 s. Each measurement is the mean of at least five independent determinations. Spectra processing and analysis were performed with the software MestReNova.[24] The fractional STD effect was calculated by $(I_0\text{-}I)/ I_0$, in which $(I_0\text{-}I)$ is the peak intensity in the STD spectrum and I_0 is the peak intensity in the off-resonance spectrum.[25]

Acknowledgements

This work was supported by the Italian Cystic Fibrosis Research Foundation (Grant FFC#10/2008) with contributions from Ferrara and Bologna FFC Delegations, Comacchio Supporting Group, and Associazione Trentina FC. S.S. is a recipient of a fellowship from the Italian Cystic Fibrosis Research Foundation. We thank Prof. Jesus Jimenez-Barbero for his helpful reading of the manuscript.

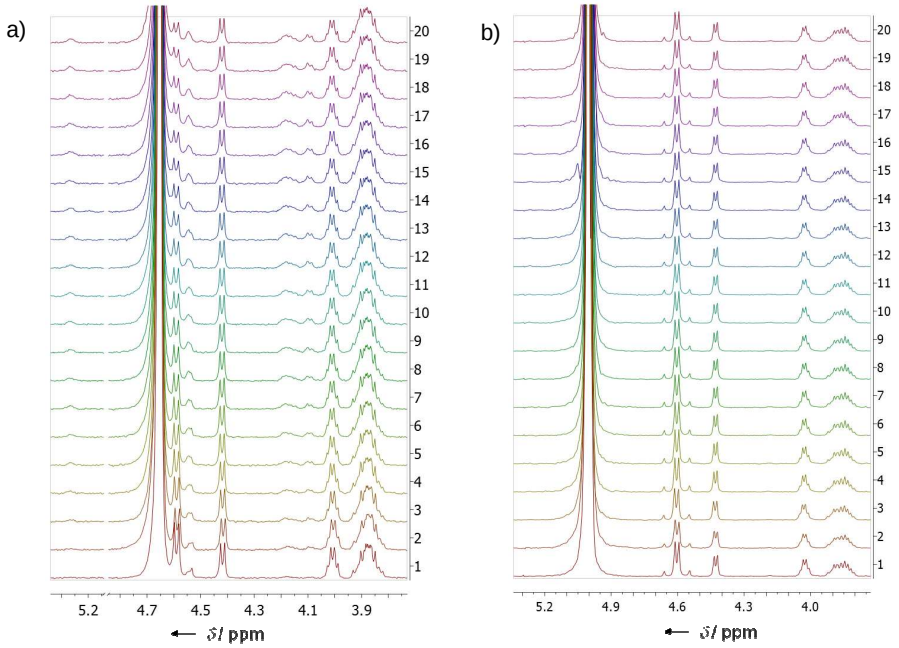
[1] C. R. Raetz, C. Whitfield, Annu. Rev. Biochem. 2002, 71, 635 –700.
[2] H. Nikaido, Microbiol. Mol. Biol. Rev. 2003, 67, 593 – 656.
[3] K. Miyake, Trends Microbiol. 2004, 12, 186 – 192.
[4] C. R. Raetz, Z. Guan, B. O. Ingram, D. A. Six, F. Song, X. Wang, J. Zhao, J. Lipid Res. 2008, 50, S103–S108.
[5] O. Holst, FEMS Microbiol. Lett. 2007, 271, 3– 113.

[6] L. Steeghs, R. den Hartog, A. den Boer, B. Zomer, P. Roholl, P. van der Ley, Nature 1998, 392, 449 – 450.
[7] C. A. Schnaitman, J. D. Klena, Microbiol. Rev. 1993, 57, 655 – 682.
[8] T. C. Meredith, P. Aggarwal, U. Mamat, B. Lindner, R. W. Woodard, ACS Chem. Biol. 2006, 1, 33– 42.
[9] G. Klein, B. Lindner, W. Brabetz, H. Brade, S. Raina, J. Biol. Chem. 2009, 284, 15369 –15389.
[10] T. C. Meredith, R. W. Woodard, J. Bacteriol. 2005, 187, 6936 –6942.
[11] T. C. Meredith, R. W. Woodard, J. Biol. Chem. 2003, 278, 32771 – 32777.
[12] P. Sperandeo, C. Pozzi, G. Deho, A. Polissi, Res. Microbiol. 2006, 157, 547 –558.
[13] A. Bateman, Trends Biochem. Sci. 1999, 24, 94–95.
[14] S. Sommaruga, L. De Gioia, P. Tortora, A. Polissi, Biochem. Biophys. Res. Commun. 2009, 388–390, 222 –227.
[15] B. Meyer, T. Peters, Angew. Chem. 2003, 115, 890 –918; Angew. Chem. Int. Ed. 2003, 42, 864 –890.
[16] T. Biet, T. Peters, Angew. Chem. 2001, 113, 4320 –4323; Angew. Chem. Int. Ed. 2001, 40, 4189 – 4192.
[17] O. Berteau, C. Sandstrçm, J. Bielicki, D. S. Anson, L. Kenne, J. Am. Chem. Soc. 2003, 125, 15296 –15297.
[18] Y. Yuan, X. Wen, D. A. R. Sanders, B. M. Pinto, Biochemistry 2005, 44, 14080 – 14089.
[19] L. Brecker, G. D. Straganz, C. E. Tyl, W. Steiner, B. Nidetzky, J. Mol. Catal. B 2006, 42, 85– 89.
[20] Y. Yuan, X. Wen, D. A. R. Sanders, B. M. Pinto, J. Am. Chem. Soc. 2008, 130, 3157 –3168.
[21] J. Yan, A. D. Kline, H. Mo, M. J. Shapiro, E. R. Zartler, J. Magn. Reson. 2003, 163, 270 –276.
[22] A. Kohen, A. Jackob, T. Baasov, Eur. J. Biochem. 1992, 208 , 443 – 449. [23] M. Mayer, B. Meyer, J. Am. Chem. Soc. 2001, 123, 6108 – 6117.
[24] http://mestrelab.com/.
[25] Y. Yuan, X. Wen, D. A. R. Sanders, B. M. Pinto, Biochemistry 2005, 44, 14080 – 14089.

Supporting Information

Monitoring of KdsD-catalyzed isomerisation by ^{1}H-NMR

The kinetic of the enzyme-catalyzed isomerisation was monitored by ^{1}H-NMR to find the experimental conditions (buffer composition, pH and temperature) more suitable for STD experiments. Examples concerning the kinetic of isomerisation in PBS, pH 7.3, NaCl 300mM, 37°C or 5°C are reported for comparison in Figure 1S.



d)

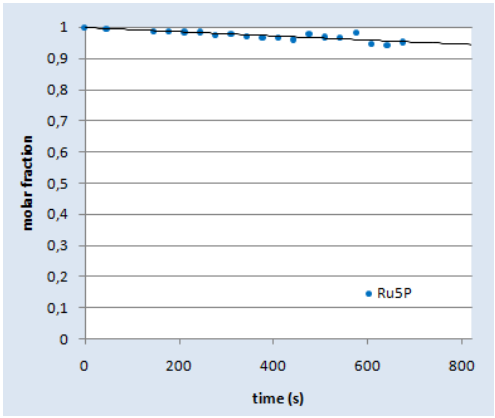


Figure 1S. a) ^{1}H NMR spectra of Ru5P 3mM + KdsD 500nM, 37°C. b) ^{1}H NMR spectra of Ru5P 3mM + KdsD 500nM, 5°C. Number of scans (NS)=8; spectrum 1 recorded at t=0s, spectra 2-20 recorded in 33s time intervals over 673s. c) Decrease of Ru5P molar fraction vs time at 37°C. d) Decrease of Ru5P molar fraction vs time at 5°C.

In order to prove that KdsD secondary structure was retained after the temperature shift from 37°C (physiological temperature) to 5°C the CD spectra of the protein recorded at the two temperatures were compared, resulting identical as shown in Figure 2S.

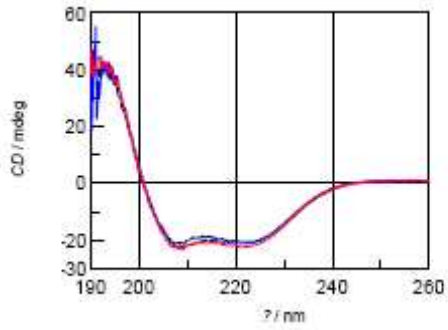


Figure 2S. CD spectra of 10 µM KdsD in PBS, pH 7.3, at 5°C (black line), at 37°C after shift from 5°C (blue line) and at 5°C after shift from 37°C (red line).

Competitive STD spectra with R5P

R5P competes with A5P and Ru5P for KdsD binding site, as proved by competitive STD experiments (Figure 3S). R5P was added to a mixture of KdsD and A5P and Ru5P at the thermodynamic equilibrium (total substrate concentration 10mM, A5P:Ru5P 7:3) at different concentrations; for each R5P concentration a STD spectrum of the mixture was acquired.

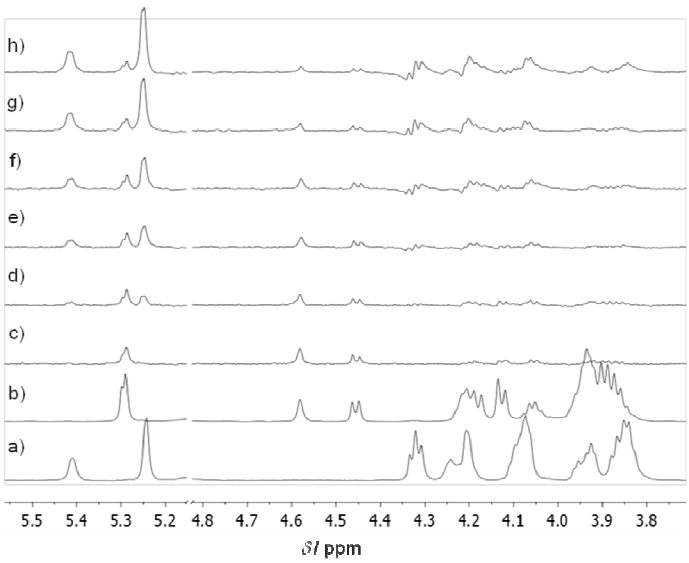


Figure 3S. a) ^{1}H-NMR spectrum of R5P 3.5 mM, NS=16. b) ^{1}H-NMR spectrum of KdsD : natural substrate mixture at equilibrium; A5P, Ru5P and KdsD concentrations were respectively 7mM, 3mM and 30 µM; NS=16. c-h) STD spectra of KdsD : substrate mixture at equilibrium added of R5P at different concentrations (c), 0mM; (d), 3mM; (e), 7mM; (f), 10mM; (g), 20mM; (h), 30mM. For all the STD experiments: NS=768, on-resonance frequency=7.9 ppm, off-resonance frequency=40 ppm, protein saturation time=2s. All the samples were dissolved in PBS, pH 7.3, 5°C.

Journal of Carbohydrate Chemistry, 29:30–38, 2010

Synthesis of 3-Deoxy-D-*threo*pentofuranose 5-Phosphate, a Substrate of Arabinose 5-Phosphate Isomerase

Cristina Airoldi, Silvia Merlo, and Francesco Nicotra

3-Deoxy-D-*threo*pentofuranose 5-phosphate, a substrate of arabinose 5 phosphate isomerase, has been synthesised starting from D-arabinose. Selective protection of the hydroxyl groups at C-1, C-2, and C-5 allowed deoxygenation of position 3 by conversion into a thiocarbamate and radical reduction. Deprotection and phosphorylation of the primary hydroxyl group and final deprotection of the other hydroxyl groups afforded the desired compound.

Keywords: Deoxysugars; API; LPS biosynthesis

Introduction

Arabinose 5-phosphate isomerase (API) converts ribulose 5-phosphate into arabinose 5-phosphate in bacteria. This reaction is the first step of the biosynthesis of KDO (3-deoxy-D-manno-octulosonate), an essential component of LPS (lipopolysaccharide) structure.[1] As a consequence, API activity is fundamental for Gram-negative bacteria survival,[2] which results in it being a relevant pharmacological target for the development of new antibacterial agents.[3,4]

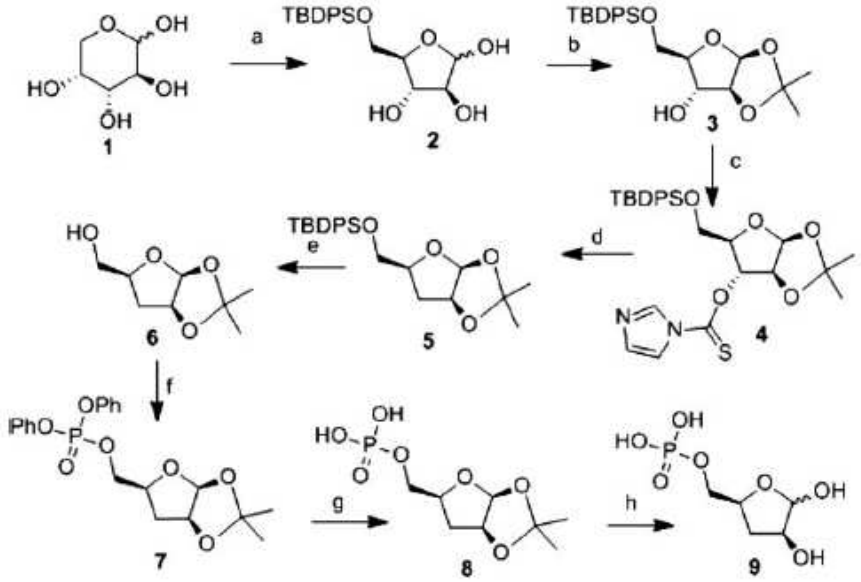
We therefore decided to investigate the structural requirements of the active site of the enzyme in order to design inhibitors with potential antibacterial activity. In particular, we studied several modifications on arabinose 5-phosphate to verify their effect on substrate recognition by the enzyme. In this context, we synthesised a 3-deoxy derivative of the natural

substrate, namely 3-deoxy-D-*threo*pentofuranose 5-phosphate, and we verified API ability to recognize and convert it into the corresponding ketose. 3-Deoxy-D-*threo*pentofuranose 5-phosphate has been already reported as a weak API inhibitor,[5] but the synthesis is not described; in this work we demonstrate that this compound is in fact a substrate of this enzyme.

Results and discussion

The synthesis of 3-deoxy-D-*threo*pentofuranose 5-phosphate was performed starting from D-arabinose **1** (Sch. 1), which was treated first with *t*-butyldiphenylsilyl chloride (TBDPSCl) in order to protect selectively the primary hydroxyl group affording compound **2**,[6] and then with acetone dimethylacetal in dichloromethane, in the presence of camphorsulphonic acid (CSA), to protect the hydroxyl groups in positions 1 and 2. Those two hydroxyl groups are the only cis-related, in the *β*-anomer, that can generate the isopropylidene cycle. The obtained 5-*O*-*t*-butyldiphenylsilyl-1,2-*O*-isopropylidene-β-D-arabinofuranose (**3**)[6,7] has the only hydroxyl group at C-3 deprotected, allowing its reduction to hydrogen, exploiting the radical displacement of a thiocarbamate. To this purpose, compound **3** was treated with thiocarbonyldiimidazole in dry dichloroethane at 80°C to afford the carbamate **4**, which was directly treated with $(Me_3Si)_3SiH$ and AIBN in toluene under reflux,[8] affording 5-*O*-*t*-butyldiphenylsilyl-3-deoxy-1,2-*O*-isopropylidene-β-D-*threo*pentofuranose **5**[6] in 69% yield over two steps. The deoxygenated compound **5** was then deprotected at the primary hydroxyl group by treatment with tetrabutylammonium fluoride (TBAF) in THF affording 3-deoxy-1,2-*O*-isopropylidene-β-D-*threo*pentofuranose **6** in 91% yield. The primary hydroxyl group of **6** was then converted into a diphenylphosphate by treatment with diphenyl chlorophosphate in dry

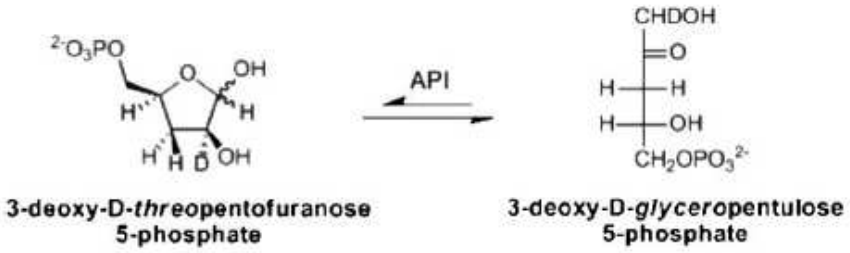
pyridine (80% yield), and the diphenylphosphate was quantitatively deprotected by treatment with PtO_2 in degassed methanol, affording 3-deoxy-1,2-*O*-isopropylidene-β-D-*threo*pentofuranose 5-phosphate **8**. Finally, hydrolysis of the isopropylidene, the last protecting group, afforded 3-deoxy-D-*threo*pentofuranose 5-phosphate **9** in quantitative yield.



Scheme 1: Reagents and conditions. (a) TBDPSCl, Py dry, 4◦C, 75% yield; (b) $Me_2C(OMe)_2$, CSA, CH_2Cl_2, 4°C, 63% yield; (c) thiocarbonyldiimidazole, CH_3CH_2Cl, rfx, 80°C; (d) $(Me_3Si)SiH$, AIBN, toluene dry, rfx, 110°C, 69% yield over two steps; (e) TBAF, THF, 91% yield; (f) diphenyl chlorophosphate, Py dry, 80% yield; (g) H_2, PtO_2, MeOH, quantitative yield; (h) H_2O, quantitative yield.

STD NMR experiments[9] performed with *Escherichia coli* API[10] showed that 3-deoxy-D-*threo*pentofuranose 5-phosphate **9** is accepted in the active siteof the enzyme, and it is isomerized to the corresponding ketose sugar (Sch. 2, Fig. 1). According to our data, positions 1 and 5 of both the deoxy substrates are fundamental for recognition and binding processes and their affinity for API is lower with respect to the natural substrate. In addition, API-specific activity for arabinose 5-phosphate and for compound **9** was

determinated by monitoring the enzymatic reaction by ^{1}H-NMR. The calculated values are $2.306\pm0.467\times10^{-4}$ and $3.295\pm0.165\times10^{-5}$, respectively. These results indicate that 3-OH deletion reduces the binding affinity, even if it does not inhibit isomerisation. This observation can be exploited in order to design and synthesize new API inhibitors as potential antibacterial drugs.



Scheme 2: The reversible isomerization of 3-deoxy-D-*threo*pentofuranose 5-phosphate to 3-deoxy-D-*glycero*pentulose 5-phosphate catalyzed by API. When the reaction is carried in D_2O, the deuteration of 3-deoxy-D-*glycero*pentulose 5-phosphate at position 1 and of 3-deoxy-D-*threo*pentofuranose 5-phosphate at position 2 occurs.

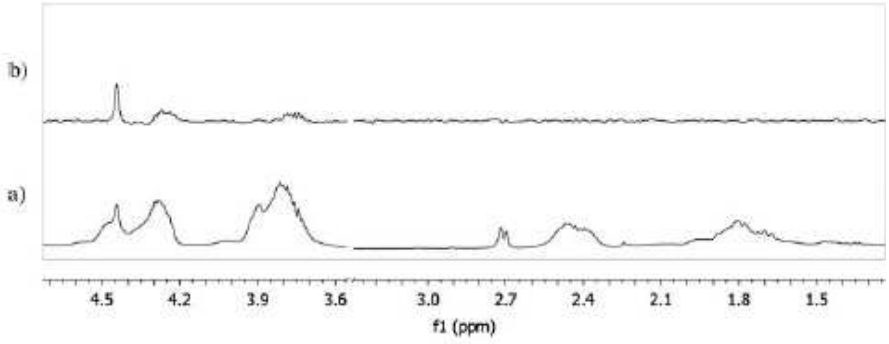


Figure 1: (a) ^{1}H NMR spectrum of API: 3-deoxy-substrate mixture at equilibrium (3-deoxy-D-*threo*pentofuranose 5-phosphate: 3-deoxy-D-*glycero*pentulose 5-phosphate NS = 16; (b) STD spectrum of API-substrate mixture at equilibrium, at an API: 3-deoxy-D-*threo*pentofuranose 5-phosphate 1:80 molar ratio and an API:3-deoxy-D-*glycero*pentulose 5-phosphate 1:120 molar ratio, protein saturation time = 2 s, NS = 2304, on-resonance frequency = 7.9 ppm, off-resonance frequency = 40 ppm. Spectra were recorded on the same sample; the sample was dissolved in PBS, pH 7.3, 5°C.

Experimental

Synthesis

General methods

All solvents were dried with molecular sieves (4 Å , Fluka) for at least 24 h prior to use. When dry conditions were required, the reactions were performed under argon. Thin layer chromatography (TLC) was performed on silica gel 60 F254 plates (Merck) with UV detection, or by using a developing solution of concd. H2SO4/E1OH/H2O (10:45:45), followed by heating at 180◦C. Flash column chromatography was performed on silica gel 230–400 mesh (Merck). Mixtures of petroleum ether (boiling range 40–60°C) and ethyl acetate were used as eluents. ^{1}H, ^{13}C, ^{1}H-^{1}H COSY, and ^{1}H-^{13}C HSQC NMR spectra were recorded with a Varian 400 MHzMercury instrument at 25°C unless otherwise stated. Chemical shifts are reported in ppm downfield from TMS as internal standard; mass spectra were recorded with a Q-TRAP 2000 instrument (Applied Biosystem). Optical rotations were measured at ambient temperature, using the sodium-D line, with a POLAX-2L electronic polarimeter (ATAGO).

5-O-t-Butyldiphenylsilyl-D-*arabinofuranose (2)*

To a stirred suspension of D-arabinose (100 mg, 0.666 mmol) in dry Py (2 mL), TBDPSC1 (188 μL, 0.733 mmol) was added dropwise at 0°C under argon atmosphere. Afer 10 min the reaction was stirred at 4°C overnight. Water was added to quench the reaction and the mixture was stirred for 15 min. The mixture was concentrated under reduced pressure. The product was purified by flash column chromatography (5:5, petroleum ether;EtOAc) affording **2** (194 g, 75% yield) as a yellow oil (mixture of α-and β-anomers). ^{1}H NMR (400 MHz, $CDCl_3$) δ = 7.83–7.56 (m, 4H, Ar), 7.56–7.27 (m, 6H, Ar), 5.41 (s, 1H, H1), 5.29 (d, 1H, J = 3.8, H1), 4.39–3.57 (m, 5H), 1.21–0.97 (s, 9H, 3 × CH3). ^{13}C NMR (101 MHz, $CDCl_3$) δ 136.05, 135.87,

135.83, 132.25, 132.05, 130.43, 130.34, 130.26, 128.25, 128.20, 128.18, 128.01, 103.49, 96.72, 86.74, 82.92, 79.60, 78.03, 77.86, 76.28, 65.01, 64.39, 60.86, 27.10, 26.95, 19.29. MS: m/z = 411.4 [M+Na]+; m/z = 427.4 $[M+K]^+$ Anal. calcd. for C21H28O5Si: C, 64,92; H, 7.26; O, 20.59; Si, 7.23; found C, 64.95; H, 7.24; O, 20.58; Si, 7.24.

5-O-t-Butyldiphenylsilyl-1,2-O-isopropylidene-β- D-*arabinofuranose (3)*

To a stirred solution of **2** (50 mg, 0.129 mmol) in dry CH_2Cl_2 (2.5 mL), 2,2-dimethoxypropane (24 μL, 0.194 mmol) and camphorsulphonic acid (catalytic amount) were added at 4°C under argon atmosphere. The reactionmixture was stirred overnight at 4°C. After 14 h, Et_3N was added to neutralize the reaction and the mixture was stirred for 15 min. The mixture was concentrated under reduced pressure. The product was purified by flash column chromatography (8:2, petroleum ether:EtOAc) giving **3** (35 mg, 63% yield) as a yellow oil.

^{1}H NMR (400 MHz, $CDCl_3$) δ = 7.78–7.59 (m, 4H, Ar), 7.50–7.31 (m, 6H, Ar), 5.89 (d1H, *J* = 4.0, H1), 4.55 (d, 1H, *J* = 4.0, H2), 4.43 (d, 1H, *J* = 1.9, H3), 4.19–3.98 (m, 1H, H4), 3.98–3.72 (m, 2H, 2 × H5), 1.34 (s, 3H, CH3), 1.29 (s, 3H, 3H, CH3), 1.10 (s, 9H, 3 × CH3). ^{13}C NMR (101 MHz, $CDCl_3$) δ 135.82, 133.35, 130.02, 128.01, 112.76, 105.81, 87.66, 87.29, 87.23, 76.49, 63.86, 27.09, 26.35, 19,45. $[\alpha]_D^{20}$ = –12.1 (c = 1, CHCl3); MS: m/z = 451.4 [M+Na]+; m/z—467.4 $[M + K]^+$ Anal. calcd. for $C_{24}H_{32}O_5Si$: C, 67.26; H, 7.53; O, 18.67; Si, 6.55; found: C, 67.22; H, 7.54; O, 18.66; Si, 6.57.

5-O-t-Butyldiphenylsilyl-3-O-*(imidazolylthiocarbonyl)-1,2–O-isopropylidene*-β- D-*threopentofuranose (4)*

To a stirred solution of **3** (65 mg, 0.151 mmol) in dry $CH_3CH_2Cl_2$ (2 mL), 1,1'-thiocarbonyldiimidazole (35 mg, 0.197 mmol) was added under argon atmosphere. The reaction mixture was heated at reflux (80°C). After 3 h, the reaction mixture was cooled to r.t. and the mixture was concentrated under reduced pressure. The residue was dissolved in CH_2Cl_2 and the organic layer was washed with water.

5-O-t-Butyldiphenylsilyl-3-deoxy-1,2-O-isopyropylidene-β-D-*threopentofuranose (5)*

To a stirred solution of crude **4** in dry toluene (2 mL), tris(trimethylsilyl) silane (33 μL, 0.196 mmol) and α,α'-azobisbutyronitrile (9 mg, 0.055 mmol) were added and the resulting solution was heated and stirred at 110°C under argon atmosphere and refluxed. After 20 min, the mixture was cooled to r.t. and the mixture was concentrated under reduced pressure. The product was purified by flash column chromatography (9.5:0.5, petroleum ether:EtOAc) affording **5** (42 mg, 69% yield over 2 steps) as a yellow oil.

^{1}H NMR (400 MHz, $CDCl_3$) δ–7.72–7.63 (m, 4H, Ar), 7.46–7.32 (m, 6H, Ar), 5.79 (d, 1H, *J* = 3.9, H1), 4.79–4.69 (m, 1H, H2), 4.37–4.23 (m, 1H, H4), 3.91–3.73 (m, 2H, 2 × H5), 2.27 (dd, 1H, *J* = 14.2, 2.4, H3a), 2.15 (ddd, 1H, *J* = 14.3, 8.1, 6.2, H3b), 1.33 (s, 3H, CH_3), 1.28 (s, 4H, CH_3), 1.06 (s, 9H, 3 × CH_3).
^{13}C NMR (101 MHz, $CDCl_3$) δ 135.86, 133.74, 129.86, 127.90, 112.33, 106.87, 81.49, 80.96, 66.06, 33.78, 27.20, 27.06, 26.14, 19.46, $[\alpha]_D^{20}$ = –10 (c = 1, $CHCl_3$); MS: m/z = 413.7 [M+H]+, m/z = 435.7 $[M+Na]^+$. Anal. calcd. for $C_{24}H_{32}O_4Si$: C, 69,86; H,7.82; O, 15.51; Si, 6,81; found: C, 69.87; H, 7.84; O, 15.49; Si, 6.80.

3-Deoxy 1,2-O-isopropylidene-β-D-*threopentofuranose (6)*

To a stirred solution of **5** in dry THF (2 mL), tetrabutylammonium fluoride 1M in THF (121 μL, 0.121 mmol) was added at r t under argon atmosphere, After 1 h, silica was added and the mixture was conecentrated under reduced pressure. The product was purified by flash column chromatography (9.5:0.5, petroleum ether: $(CH_3)_2CO$) giving **6** (19.3 mg 91% yield) as a yellow oil.

^{1}H NMR (400 MHz, $CDCl_3$) δ = 5.81 (d, 1H, *J* = 3.8, H1), 4.75 (m, 1H, H2), 4.31 (m, 1H, H4), 3.81 (dd, 1H *J*–11.5, 8.0, IH, H5a), 3.60 (dd, 1H, *J* = 11.5, 3.9, H5b), 2.19 (ddd, 1H *J* = 14.7, 8.6, 6.3, H3a), 1.98 (dd, 1H *J* = 14.2, 3.3, H3b), 1.54 (s, 3H, CH_3), 1.31 (s, 3H, CH_3). ^{13}C NMR (101 MHz, $CDCl_3$) δ 112.35, 106.67, 81.95, 80.91, 65.23, 33.39, 27.21, 26.13. $[\alpha]_D^{20}$ = +4.9 (c = 1, $CHCl_3$);

MS: m/z = 197.2 $[M+Na]^+$. Anal. calcd. for $C_8H_{14}O_4$; C, 55.16; H, 8.10; O, 36.74; found: C, 55.14; H, 8.11; O, 36.70.

*3-Deoxy-1,2-*O*-isopropylidene-*β-D-*threopentofuranose-5-diphenylphosphate (7)*

To a stirred solution of **6** (220 mg, 1.26 mmol) in dry Py (13 mL), diphenylphosphochloridate (24 μL, 0.194 mmol) was added at 0°C under argon atmosphere. After 1 h, MeOH was added to quench the reaction and the mixture was stirred for 15 min. The mixture was concentrated under reduced pressure. The product was purified by flash column chromatography (6:4, petroleum ether:EtOAc) affording **7** (35 mg, 80% yield) as a white solid.

^{1}H NMR (400 MHz, $CDCl_3$) δ = 7.41–6.93 (m, 10H, Ar), 5.75 (d, 1H, *J* = 3.7, H1), 4.64 (bt, 1H, *J* = 4.1, H2), 4.52–4.38 (m, 1H, *J* = 9.5, 7.1, H5a), 4.39–4.30 (m, 1H, *J* = 8.6, 2.4, H4), 4.30–4.21 (m, 1H, *J* = 9.6, 6.6, H5b), 2.13–1.97 (m, 2H, 2 × H3), 1.46 (s, 3H, CH_3), 1.22 (s, 3H, CH_3). ^{13}C NMR (101 MHz, $CDCl_3$) δ 150.31, 130.00, 125.58, 120.32, 112.66, 107.07, 80.57, 79.04, 70.38, 33.63, 27.17, 25.95. ^{31}P NMR (162 MHz, $CDCl_3$) δ –10.96, $[\alpha]_D^{20}$ = +5.9 (c = 1, $CHCl_3$); MS: m/z = 407.3 $[M+H]^+$; m/z = 429.3 $[M+Na]^+$; m/z = 445.3 $[M+K]^+$. Anal. calcd. for $C_{20}H_{23}O_7P$: C, 59,11; H, 5.70; O, 27.56; P, 7.62; found: C, 59.09; H, 5.71; O, 27.54; P, 7.61.

*3-Deoxy-1,2-O-isopropylidene-*β-D-*threopentofuranose 5-phosphate (8)*

Compound **7** (30 mg, 0.074 mmol) was dissolved in degassed MeOH (3 ml) and PtO_2 was added (16.8 mg). The reaction mixture was stirred under H_2 at r t for 2 h. The catalyst was removed by filtration through Celite and the filtrate was concentrated under reduced pressure, obtaining **8** (19 mg, quantitative yield) as a white solid.

^{1}H NMR (400 MHz, CD_3OD) δ = 5.77 (d, 1H, *J* = 3.0, H1), 4.84–4.67 (m, 1H, H2), 4.44–4.23 (m, IH, H4), 4.23–3.86 (m, 2H, *J* = 40.3, 2 × H5), 2.34–1.94 (m, 2H, *J* = 33.4, 17.6, 2 × H3), 1.51 (s, 3H, CH_3), 1.28 (s, 3H, CH_3). ^{13}C NMR (101 MHz, CD_3OD) δ 112.38, 106.99, 80.75, 79.73, 68.07, 33.10, 26.26, 24.98. 31P NMR (162 MHz, D2O) δ 1.07. $[\alpha]_D^{20}$=+5.9 (c = 1, H_2O); MS: m/z = 255.3 $[M+H]^+$; m/z = 277.3 $[M+Na]^+$; m/z = 293.3 $[M+K]^+$. Anal. calcd. for

$C_8H_{15}O_7P$: C, 37.80; H, 5.95; O, 44.06; P, 12.19; found: C, 37.78; H, 5.97; O, 44.02; P, 12.20.

*3-Deoxy-*D*-threopentofuranose 5-phosphate (9)*

Compound **8** was dissolved in H_2O and the solution was transferred into an NMR-tube. The reaction was monitored by ^1H NMR. After 1 h pH was neutralized by adding NaOH 1.0 M and the solution was lyophilized to give **9** in quantitative yield (mixture of α-and β-anomers).

^1H NMR (400 MHz, D_2O) δ = 5.10 (s, 1H, H1), 5.02 (d, 1H, J = 4.3, H1), 4.26 (m, 1H, H2), 4.18–3.93 (m, 1H, H4), 3.93–3.58 (m, 2H, 2 × H5), 2.38–2.06 (m, 1H, H3a), 1.66–1.35 (m, 1H, H3b). 13C NMR (101 MHz, D_2O) δ 102.23, 95.63, 77.48, 77.24, 75.21, 74.35, 70.83, 67.84, 32.84, 30.94. MS: m/z = 259.1 $[M(disodium\ salt)+H]^+$. Anal. calcd. for $C_5H_{11}O_7P$: C, 28.05; H, 5.18; O, 52.31; P, 14.47; found: C, 28.07; H, 5.19; O, 52.28; P, 14.49.

NMR experiments

Binding studies

The ligand resonances were assigned using ^1H-^1H COSY and ^1H-^{13}C HSQC NMR spectroscopy. A basic 1D-STD sequence was used with the on-resonance frequency of 7.9 ppm and the off-resonance frequency of 40 ppm. A train of Gaussian-shaped pulses of 50 ms each was employed, with a total saturation time of the protein envelope of 2 s. A $T_1\rho$ filter of 2 ms was employed to eliminate the background signals from protein. All the samples were dissolved in deuterated PBS at pH 7.3, containing NaCl (300 mM), at 5°C. Total sample volumes were 550 μL. The on-resonance and the off-resonance spectra were acquired simultaneously with the same number of scans. The STD spectrum was obtained by subtraction of the on resonance spectrum from the off-resonance spectrum. Subtraction was performed by phase cycling to minimize artifacts arising from magnet and

temperature instabilities.[11] Reference experiments of samples containing only the free tested compound itself were performed under the same experimental conditions to verify true ligand binding. The effects observed in the presence of the protein were due to true saturation transfer, as no signal was present in the STD spectra obtained in the reference experiments, except residues from HDO, indicating artifacts from the subtraction of compound signals to be negligible.

Enzymatic assays

API-specific activity was determined as increase of ribulose 5-phosphate H3 or of 3-deoxy-D-*glycero*pentulose 5-phosphate H3 integral value vs. time per μg of API. Spectra were collected in 33 s time intervals over 673 s. Measurement is the mean of five independent determinations.

Acknowledgements

This work was supported by the italian cystic fibrosis research foundation (grant ffc#10/2008) and the italian consortium cinmpis.

References

1. Holst, O. The structures of core regions from enterobacterial lipopolysaccharides—an update. *FEMS Microbiol. Lett.* **2007**, *271*, 3–11.
2. Raetz, C.R.; Whitfield, C. Lipopolysaccharide endotoxins. *Annu. Rev. Biochem.* **2002**, *71*, 635–700.
3. Meredith, T.C.; Aggarwal, p.; Mamat, U.; Lindner, B.; Woodard, R.W. Redefining the requisite lipopolysaccharide structure in *Escherichia coli*. *ACS Chem. Biol.* **2006**, *1*, 33–42.
4. Klein, G.; Lindner, B.; Brabetz, W.; Brade, H.; Raina, S. *Escherichia coli* K-12 suppressor-free mutants lacking early glycosyltransferases and late acyltransferase. *J. Biol. Chem.* **2009**, *284*, 15369–15389.
5. Bingham, E.C.; Gragg, C.E.; Hall, W.R.; Kelsey, J.E.; Mallory, W.R.; Richardson, D.C.; Benedict, C.; Ray, P.H. Inhibition of Arabinose-5-phosphate isomerase. An approach to the inhibition of bacterial lipopolysaccharide biosynthesis. *J. Med. Chem.* **1984**, *27*, 717–726.
6. (a) Ikeda, M.; Bando, H. Preparation of enynes as intermediates for vitamin D3 derivatives. *Jpn. Kokai Tokkyo Koho* **1995**, 12P. (b) Sunagawa, J.; Uenishi, Y.; Kususe, N.; Ito, Y.; Ishiwatari, A.; Akao, H.; Yano, I.; Fujita, Y. Arabinomycolate compounds,

their manufacture from cell wall skeleton components of Mycobacterium bovis BCG, and their uses as antitumor agents, etc. *Jpn. Kokai Tokkyo Koho* **2006**, 44pp.
7. (a) Martin, O.R.; Rao, S.P.; El-Shenawy, H.A.; Kurz, K.G.; Cutler, A.B. Unsual 1,5-hydride shifts in lewis acid mediated reactions of benzylated sugars. Synthesis of 3-alkylisochroman derivatives. *J. Org. Chem.* **1988**, *53*, 3287–3292; (b) Fleet, G.W.J.; Mathews, C.J.; Seijas, J.A.; Vazquez Tato, M.P.; Brown, D.J. Chiral quinuclidines (1-azabicyclo[2.2.2]octanes) from sugars: synthesis of (3S,5S)-quinuclidine-3,5-diol from D-arabinose. *J. Chem. Soc. Perkin Trans 1* **1989**, *5*, 1065–1066; (c) Elliott, R.P.; Fleet, G.W.J.; Vogt, K.; Wilson, F.X.; Wang, Y.; Witty, D.R.; Storer, R.; Myers, P.L.; Wallis, C.J. Attempted ring contraction of α-triflates of 3-azido- and 3-fluoro-γ-lactiones to oxetanes. *Tetrahedron Asymm.* **1990**, *1*, 715–718; (d) Pakulski, Z.; Zamojski, A. The reaction of acctyliron [(η5-C5H5)Fe(CO)(PPh3)(COCH3)] with sugar aldehydes. New synthesis of deoxy sugars. *Tetrahedron* **1995**, *51*, 871–908; (e) Vazquez-Tato, M.P.; Seijas, J.A.; Fleet, G.W.J.; Mathews, C.J.; Hemmings, P.R.; Brown, D. Synthesis of (3S,5S)-quinuclidine- 3,5-diol and of [3S-(3α,3aα, 7aα)]-octahydro-2-furo[2,3 c]pyridinol from D-arabinose. *Tetrahedron* **1995**, *51*, 959–974; (f) Mereyala, H.B.; Hotha, S.; Gurjar, M.K. Synthesis of pentaarabinofuranosyl structure motif A of Mycobacterium tuberculosis. *Chem. Comm.* **1998**, *6*, 685–686; (g) Hoeffler, J.-F.; Grosdemange-Billiard, C.; Rohmer, M. Elucidation of the 2-C-methyl-D-erythritol 4-phosphate pathway for isoprenoid biosynthesis: straightforward syntheses of enantiopure 1-deoxy-D-xylulose from pentose derivatives. *Tetrahedron Lett.* **2001**, *42*, 3065–3067; (h) Sorensen, M. H.; Nielsen, C.; Nielsen, P. Synthesis of a bicyclic analogue of AZT restricted in an unusual 04_-Endo conformation. *J. Org. Chem.* **2001**, *66*, 4878–4886; (i) Sharma, G.V.M.; Reddy, V. G.; Radha Krishna, P.; Sankar, A.R.; Kunwar, A.C. Synthesis of spiro carbon linked disaccharides from Dglucose, D- and L-arabinose. *Tetrahedron* **2002**, *58*, 3801–3812; (l) Houseknecht, J.B.; Lowary, T.L. Oligofuranosides containing conformationally restricted residues: synthesis and conformational analysis. *J. Org. Chem.* **2002**, *67*, 4150–4164; (m) Larsen, C.H.; Ridgway, B.H.; Shaw, J.T.; Smith, D.M.; Woerpel, K.A. Stereoselective C-glycosylation reactions of ribose derivatives: electronic effects of five-membered ring oxocarbenium ions. *J. Am. Chem. Soc.* **2005**, *127*, 10879–10884; (n) Khasnobis, S.; Zhang, J.; Angala, S.K.; Amin, A.G.; McNeil, M.R.; Crick, D.C.; Chatterjee, D. Characterization of a specific arabinosyltransferase activity involved in mycobacterial arabinan biosynthesis. *Chem. Biol.* **2006**, *13*, 787–795; (o) Chevallier, O.P.; Migaud, M.E. Investigation of acetyl migrations in furanosides. *Beilstein J. Org. Chem.* **2006**, *2*(14), 1–6; (p) Doboszewski, B.; Herdewijn, P. Carbohydrate-based approach to four enantiomerically pure 2-naphthylmethyl 3-hydroxy-2-methylbutanoates. *Tetrahedron Lett.* **2008**, *49*, 1331–1335; (q) Sattigeri, V.J.; Palle, V.P.; Khera, M.K.; Reddy, R.; Tiwari, M.K.; Soni, A.; Abdul Rauf, A.R.; Joseph, S.; Musib, A.; Dastidar, S.G.; Srivastava, Punit Kumar. Preparation of uronic acid derivatives as metalloproteinase inhibitors. *PCT Int. Appl.* **2008**, 183p. (r) Doboszewski, B.; Herdewijn, P. Carbohydrate chiral-pool approach to four enantiomerically pure 2-naphthylmethyl 3-hydroxy-2-methylbutanoates. *Terahedron***2008**, *64*, 5551–5562; (s)Muraoka, O.; Tanabe, G. Preparation of cyclic sulfonium salt as glycosidase inhibitor. *PCT Int. Appl.* **2008**, 55p.

8. Mathe, C.; Imbach, J.-L.; Gosselin, G. 1,2-Di-*O*-acetyl-5-*O*-benzoyl-3-deoxy-L*erythro*- pentofuranose, a convenient precursor for the stereospecific synthesis of nucleoside analogues with the unnatural *β*-L-configuration. *Carbohydr. Res.* **2000**, *323*, 226–229.
9. Meyer, B.; Peters, T. NMR spectroscopy techniques for screening and identifying ligand binding to protein receptors. *Angew. Chem. Int. Ed.* **2003**, *42*, 864–890.
10. Meredith, T.C.;Woodard, R.W. *Escherichia coli* YrbH Is a D-arabinose 5-phosphate isomerase. *J. Biol. Chem.* **2003**, *278*, 32771–32777.
11. Mayer, M.; Meyer, B. 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*, 6108–6117.

ChemBioChem 2011, 12, 719 – 727

Targeting Bacterial Membranes: Identification of *Pseudomonas aeruginosa* D-Arabinose-5P Isomerase and NMR Characterisation of its Substrate Recognition and Binding Properties

Cristina Airoldi, Silvia Sommaruga, Silvia Merlo, Paola Sperandeo, Laura Cipolla, Alessandra Polissi, and Francesco Nicotra

The identification and characterisation of *Pseudomonas aeruginosa* KdsD (Pa-KdsD), a D-arabinose-5P isomerase involved in the biosynthesis of 3-deoxy-D-manno-oct-2-ulosonic acid and thus of lipopolysaccharide (LPS), are reported. We have demonstrated that KdsD is essential for *P. aeruginosa* survival and thus represents a key target for the development of novel antibacterial drugs. The key amino acid residues for protein activity have been identified. The structural requirements for substrate recognition and binding have been characterised for the wildtype protein, and the effect of mutations of the key residues on catalytic activity and binding have been evaluated by saturation transfer difference (STD) NMR spectroscopy. Our data provide important structural information for the rational design of new KdsD inhibitors as potential antibacterial drugs.

Introduction

Approximately 10% of all hospital-acquired infections are due to *Pseudomonas aeruginosa*, a Gram-negative opportunistic human pathogen. This bacterium infects immunosuppressed patients and causes high fatality rates in patients hospitalized with cancer, cystic fibrosis and burns.[1] *P. aeruginosa* is unique because it possesses intrinsic resistance to a variety

of antimicrobial agents as a consequence both of its low outer membrane permeability[2] and of the expression of several drugs efflux pumps.[3] Lipopolysaccharide (LPS) represents a major virulence factor of these inflammation-eliciting bacteria and is thus a primary target for effective immunity.[4] LPS, an essential glycolipid located in the outer membranes (OMs) of Gram-negative bacteria, is composed of three distinct regions: lipid A, core oligosaccharide and *O*-antigen. *P. aeruginosa* produces two forms of *O*-antigen: the homopolymeric A-band and the heteropolymeric B-band.[5, 6] The genetic clusters associated with the biosynthesis of A-band and B-band *O*-antigen in *P. aeruginosa* have been well characterised,[7] whereas relatively little is known about the biosynthesis of core oligosaccharide. As has also been observed for other Gram-negative bacteria, the *P. aeruginosa* core can be subdivided into the inner and the outer core; the former is more conserved and contains two 3-deoxy-D-manno-oct-2-ulosonic acid (Kdo) residues.[8]

In *Escherichia coli* the biosynthesis of Kdo involves four sequentially acting enzymes (Scheme 1); the first is a D-arabinose 5-phosphate isomerase (API), converting D-ribulose 5-phosphate (Ru5P) into D-arabinose 5-phosphate (A5P). In *E. coli* the API function is redundant because it is coded by the two paralogous genes *kdsD* and *gutQ*, which share significant sequence similarity and can be substituted by each other.[9, 10,11] API is indeed an essential function: in *E. coli* the Kdo2-lipid A moiety represents the minimum structure indispensable for growth,[12] inactivation of both *kdsD* and *gutQ* genes leads to nonviable cells,[10,11] and interruption of Kdo biosynthesis leads to the arrest of cell growth.[13] *P. aeruginosa* has more stringent requirements in terms of the sugars of the full inner core and at least part of the outer core. These features, in addition to lipid A, are needed for microorganism viability.[14, 15] This makes Kdo biosynthesis an

ideal target for new antibiotic activity discovery. *E. coli* KdsD is a 328-residue enzyme, active as a homo-tetrameric protein.[9, 10] Each of the four subunits contains two distinct domains: an N-terminal sugar isomerase (SIS) domain (residues 1–200), followed by a pair of cystathionine-b-synthase (CBS) domains (residues 213–325) of unknown function. The three-dimensional structure of *E. coli* KdsD has recently been predicted by homology modelling and a few amino acid residues relevant for enzyme activity have been identified.[16] Moreover, with the aid of saturation transfer difference (STD) NMR spectroscopy the structural requirements necessary for *E. coli* KdsD substrate recognition and binding have been identified.[17] The reaction catalysed by KdsD is an aldose–ketose isomerisation in which a hydrogen atom is transferred between the C1 and C2 atoms of the substrate. The carbon-bound hydrogen can be transferred by two possible mechanisms: a hydride shift or a proton transfer through an acid–base mechanism involving a cis-enediol intermediate.[18] In the xylose isomerase family the hydride shift mechanism appears to operate.[19] In the rabbit phosphoglucose isomerase PGI, which catalyses the conversion of glucose 6-phosphate into fructose 6-phosphate, an active site glutamate (Glu357) is the base catalyst for isomerisation and His388 is the acid catalyst for ring opening.[20, 21] In contrast, the isomerisation reaction catalysed by glucosamine 6-phosphate synthase from *E. coli*, which converts fructose 6-phosphate into glucosamine 6-phosphate, is likely to involve the formation of a Schiff base with Lys603, a ring-opening step catalysed by His504 and a proton transfer from C1 to C2 catalysed by Glu488.[22] The proposed mechanism for the action of *E. coli* GmhA involves an enediol switch, with Glu65 and His180 acting as an acid and a base, respectively, to promote the conversion of sedoheptulose 7 phosphate into D-*glycero*-D-*manno*-heptopyranose 7-phosphate.[23] Very

recently, however, a modified mechanism has been proposed for *Burkolderia pseudomallei* GmhA: this enzyme contains in the active site a Zn^{2+} ion that orients the side chains of Glu68 and Gln175, acting as the base and the acid respectively, to promote the isomerisation reaction.[24] In this work we report the characterisation of the kdsD homologue from *P. aeruginosa* and provide evidence that the gene is essential for growth. Moreover, on the basis of the recently reported *E. coli* KdsD structural data, we identify key amino acid residues for the activity of the *P. aeruginosa* homologue. Finally, with the aid of saturation transfer difference (STD) NMR spectroscopy, the structural requirements necessary for substrate recognition and binding in the wild-type protein have been characterized and the effect of the key-residue mutations on catalytic activity and binding have been evaluated. Collectively these data provide important structural information for the rational design of new KdsD inhibitors as potential antibiotic drugs.

Results and Discussion

The P. aeruginosa kdsD gene is essential

In *E. coli*, *kdsD* (Ec-*kdsD*) belongs to the *yrbG-lptB* locus, which has been extensively characterised and found to contain essential genes (*lptA, lptB* and *lptC*) implicated in LPS transport to the OM.[25] Analysis of the annotated *P. aeruginosa* PAO1 genome sequence identified a single putative ORF (PA4457) exhibiting similarity with the *kdsD-gutQ* family of sugar isomerases. PA4457 is located in a region displaying a very similar organisation to the corresponding cluster in *E. coli* (Figure S1 in the Supporting Information). These observations suggest that PA4457 (Pa-*kdsD*) might code for an essential function. To test whether Pa-*kdsD* is essential, we attempted its inactivation by single crossover disruption. However,

because Pa-*kdsD* is located just upstream of putative essential genes (Figure S1), we engineered the pDM4 vector (pDM4-*araBp*-*kdsD*) so that, upon integration, the arabinose-inducible promoter *araBp* drives the expression of the genes downstream of the disruption to avoid the possibility that the insertion of the cassette might exert a polar effect on downstream genes (see the Experimental Section).

We selected algR (pDM4-*araBp*-*algR*), a nonessential gene of *P. aeruginosa*, as a positive control for gene inactivation.[26] No *P. aeruginosa* Cm^R recombinants with disrupted *kdsD* gene could be obtained upon conjugation between *E. coli* harbouring the pDM4-*araBp*-*kdsD* construct, whereas *algR* disruption was obtained in crosses with pDM4-*araBp*-*algR*, as confirmed by PCR analysis (Table 1).

Table 1. Single-crossover gene disruption of conjugation efficiency.

Recipient strain	Target gene	No. of productive conjugations[a,b]	No. of integration events[c]
P. aeruginosa PAO1	*algR*	7/10	14/14
P. aeruginosa PAO1	*kdsD*	0/10	0
P. aeruginosa PAO1/Ec-*kdsD*	*algR*	6/6	10/12
P. aeruginosa PAO1/Ec-*kdsD*	*kdsD*	5/7	5/8

[a] Selection of exconjugants was performed on PIA (*Pseudomonas* isolation agar) supplemented with chloramphenicol (85 mg mL^{-1}). In Pa-*kdsD* inactivation experiments with *P. aeruginosa* PAO1/Ec-*kdsD*, arabinose (0.2%) and carbenicillin (200 mg mL^{-1}) were added. [b] The numbers of exconjugants in productive conjugations range from one to 10. [c] The integration within the target gene in the PAO1 genome was confirmed by PCR; two clones from each independent conjugation experiment were analysed.

As shown in Table 2, Pa-KdsD can complement the *E. coli* BB-8 conditional mutant[11] in which *kdsD* is under arabinose control and the *gutQ* paralogous copy is inactivated, indicating that Ec-*kdsD* and Pa-*kdsD* are functionally homologous and can substitute for each other. We therefore attempted to disrupt the Pa-*kdsD* gene in a recipient carrying the wild-type Ec-*kdsD* expressed from a plasmid. Under these conditions, Cm^R

recombinants were obtained, suggesting that disruption of the chromosomal Pa-*kdsD* can take place when another functional copy of the gene is provided in trans (Table 1). Overall, these data indicate that Pa-*kdsD* is essential for *P. aeruginosa* viability.

Table 2. Complementation analysis and specific activity of *P. aeruginosa* wild-type and mutant KdsD proteins.

Complementing plasmid	Codon change	Growth conditions[a] +*Ara*	−*Ara*	Protein	Specific activity[b] [U mg^{-1}]	Residual activity
pET14	–	+	–	–	–	–
pET14-*kdsD*	–	+	+	KdsD	29.50 ± 0.80	100%
pET14-*kdsD* Lys56Ala	AAA→GCA	+	–[c]	KdsD-Lys56Ala	0.092 ± 0.007	0.31%
pET14-*kdsD* His85Ala	CAT→GCT	+	–[c]	KdsD-His85Ala	b.d.[d]	n.a.[e]
pET14-*kdsD* His190Ala	CAC→GCC	+	–[c]	KdsDHis190Ala	b.d.[d]	n.a.[e]

[a] LD-ampicillin agar plates either supplemented or not supplemented with arabinose (Ara) for induction of BB-8 chromosomal *kdsD* gene. [b] Specific activity was determined as reported in the Experimental Section. Each measurement is the mean of at least five independent determinations. Standard deviations never exceeded 10%. [c] Efficiency of plating $\leq 10^{-4}$. [d] b.d. = below detection. [e] n.a. = not applicable.

Over-expression, purification and activity of Pa-KdsD

The recombinant Pa-KdsD protein was over-expressed by use of the pQE31 vector fused to a His6-tag at the N-terminal end. Optimal conditions for the maximum ratio of recombinant to cellular protein were achieved by induction for 4 h with isopropyl-β-D-thiogalactopyranoside (IPTG, 0.5 mM) at 37°C. Efficient purification to homogeneity was obtained by means of single step purification by Ni-chelate affinity chromatography. The subunit molecular weight was 34 kDa (MW_{calcd}=34067), as determined by SDS-PAGE electrophoresis under reducing conditions (data not shown). The specific activity was determined with A5P as a substrate, as described in the Experimental Section (Table 2).

NMR characterisation of Pa-KdsD's substrate recognition and binding requirements

Pa-KdsD's substrate recognition and binding requirements were successfully elucidated by exploiting the same experimental protocol as previously adopted to study the homologous enzyme Ec-KdsD.[17] This

approach is based on the use of saturation transfer difference (STD) NMR spectroscopy, extensively described as an efficient method of investigating protein– ligand interactions.[27–33] In particular, in analogy with the approach used for Ec-KdsD, we have characterised the binding between Pa-KdsD and its natural substrates, and also a number of substrate analogues. Firstly, we verified that the experimental conditions (buffer composition, pH and temperature) adopted for STD experiments on Ec-KdsD were suitable for STD experiments on Pa-KdsD. Similarly to what had been described for Ec-KdsD, we found that the specific activity of Pa-KdsD decreases significantly when the temperature is shifted from 37 to 5°C. In addition, circular dichroism (CD) spectra demonstrated that the Pa-KdsD secondary structure was the same at 37 and 5°C (data not shown). We therefore performed the STD experiments by dissolving our samples in phosphate buffer saline (PBS, pH 7.3) at 5°C and indeed verified that, under the experimental conditions described, substrate interconversion does not affect the measured STD effects. Hence, in all cases in which the tested compound was a Pa-KdsD substrate— and thus was converted into the corresponding isomer— STD spectra were acquired on the mixture of the aldose and ketose sugars, after the thermodynamic equilibrium had been reached. In addition, for all the interacting compounds, six different STD spectra were acquired with total saturation times of the protein envelope of 5, 3, 2, 1.5, 1 or 0.5 s, and the corresponding STD build-up curves were derived (Figure S2).

Pa-KdsD showed substrate recognition and binding requirements similar to those of the *E. coli* homologue. The prominent role played by the protons at C-1 and C-3 in the interaction between the enzyme and the natural substrates was confirmed by STD spectra recorded on the mixture at equilibrium (A5P/Ru5P 7:3; Figure 1 A).

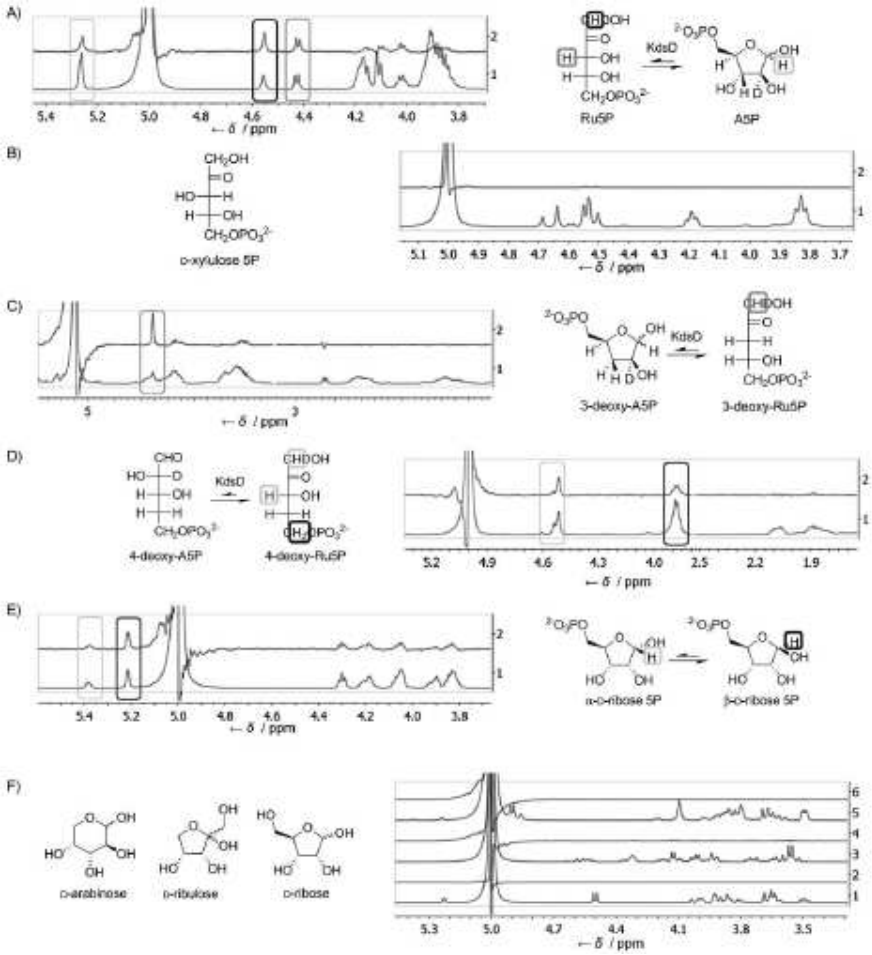


Figure 1. NMR characterisation of Pa-KdsD binding requirements. A) On the left: 1) ^{1}H NMR spectrum of Pa-KdsD/substrate mixture at equilibrium (A5P/Ru5P 7:3), and 2) STD spectrum of Pa-KdsD/substrate mixture at equilibrium, at a KdsD/A5P 1:233 molar ratio and a Pa-KdsD/Ru5P 1:100 molar ratio. On the right: the reversible isomerisation of Ru5P to A5P catalysed by Pa-KdsD. When the reaction is carried out in D_2O, deuteration of Ru5P at position 1 and of A5P at position 2 occurs. B) On the right: 1) ^{1}H NMR spectrum of Pa-KdsD/Xu5P mixture, and 2) STD spectrum of Pa-KdsD/Xu5P mixture, at a KdsD/Xu5P 1:117 molar ratio. On the left : structure of Xu5P. C) On the left: 1) ^{1}H NMR spectrum of Pa-KdsD/3-deoxy substrate mixture at equilibrium (3-deoxy-A5P/3-deoxy-Ru5P 4:6), and 2) STD spectrum of Pa-KdsD/3-deoxy-substrate mixture at equilibrium, at a KdsD/3-deoxy A5P 1:80 molar ratio and a KdsD/3-deoxy-Ru5P 1:120 molar ratio. On the right: the reversible isomerisation of 3-deoxy-A5P to 3-deoxy-Ru5P catalysed by Pa-KdsD. When the reaction is carried out in D_2O, deuteration of 3-deoxy-Ru5P at position 1 and of

3-deoxy-A5P at position 2 occurs. D) On the right: 1) ^{1}H NMR spectrum of Pa-KdsD/4-deoxy-substrate mixture at equilibrium (4-deoxyA5P/4-deoxy-Ru5P 1:9), and 2) STD spectrum of Pa-KdsD/4-deoxy-substrate mixture at equilibrium, at a KdsD/4-deoxy-A5P 1:12 molar ratio and a KdsD/4-deoxy-Ru5P 1:105 molar ratio. On the left: the reversible isomerisation of 4-deoxy-A5P to 4-deoxy-Ru5P catalysed by Pa-KdsD. When the reaction is carried out in D_2O, deuteration of 4-deoxy-Ru5P at position 1 and of 4-deoxy-A5P at position 2 occurs. E) On the left: 1) ^{1}H NMR spectrum of Pa-KdsD/R5P mixture, and 2) STD spectrum of Pa-KdsD/R5P mixture, at a KdsD/R5P 1:117 molar ratio. On the right: R5P exists in solution as a mixture of a and b anomers. F) On the right: 1) ^{1}H NMR spectrum of Pa-KdsD/D-arabinose mixture, 2) STD spectrum of Pa-KdsD/D-arabinose mixture, 3) ^{1}H NMR spectrum of Pa-KdsD/D-ribulose mixture, 4) STD spectrum of Pa-KdsD/D-ribulose mixture, 5) ^{1}H NMR spectrum of Pa-KdsD/D-ribose mixture, and 6) STD spectrum of Pa-KdsD/D-ribose mixture. For all the samples KdsD/tested compound 1:117 molar ratio. On the left: structures of D-arabinose, D-ribulose and D-ribose. All the samples were dissolved in PBS (pH 7.3, 5°C)^{1}H NMR spectra were each recorded with 16 scans. STD spectra were each recorded with 768 scans, except for the STD spectrum of the 3-deoxy-A5P/3-deoxy-Ru5P mixture, acquired with 2304 scans; protein saturation time=2 s, on-resonance frequency=7.9 ppm, off-resonance frequency= 40 ppm.

STD experiments on D-xylulose 5-phosphate (Xu5P), the D-ribulose 5-phosphate epimer at position 3, showed that Pa-KdsD does not bind to Xu5P, clearly indicating that inversion of the configuration at position 3 inhibits the recognition and binding processes (Figure 1 B).

The importance of 3-OH was further demonstrated on investigation of the ability of Pa-KdsD to convert 3-deoxy-D-*threo*pentofuranose 5-phosphate (3-deoxy-A5P)[34] into the corresponding ketose 3-deoxy-D-*glycero*-pentulose 5-phosphate (3-deoxy-Ru5P). The mixture composition at thermal equilibrium is 3-deoxy-A5P/3-deoxy-Ru5P 4:6. As in the case of Ec-KdsD, Pa-KdsD accepts 3-deoxy-A5P as substrate. However, its catalytic efficiency—the measured specific activity is (0.022±0.002) Umg^{-1} at 37°C—is about 1000 times lower than for the natural substrate isomerisation [measured specific activity (17.59±1.83) Umg^{-1} at 37°C]. Both the specific activities were determined as reported in the Experimental Section. In addition, STD intensities were significantly lower relative to the spectrum

recorded on the mixture of natural substrates, as a consequence of a decrease in affinity (Figure 1 C).

All together, these pieces of evidence indicated that 3-OH deletion does not inhibit isomerisation, but significantly affects the binding affinity.

We employed 4-deoxy-D-*threo*-pentofuranose 5-phosphate (4-deoxy-A5P) to investigate the role of 4-OH for substrate recognition and transformation. The compound was synthesized as reported in 1992 by Kohen et al.[35] As verified for Ec-KdsD, Pa-KdsD converts 4-deoxy-A5P to 4-deoxy-D-*glycero*-pentulose-5-phosphate (4-deoxy-Ru5P). The specific activity was determined as an increase in the 4-deoxy-Ru5P H4 integral value versus time per ug of KdsD. The isomerisation rate at 37°C for this compound—(12.37±2.133) Umg^{-1} at 37°C—is comparable with that measured for the natural substrate. STD experiments acquired on the mixture (4-deoxy-A5P/4-deoxy-Ru5P 1:9 at thermal equilibrium) confirmed the ability of KdsD to recognize and to bind 4-deoxy-A5P and 4-deoxy-Ru5P (Figure 1D). These data indicate that the presence of a hydroxy group at position 4 is not essential for substrate recognition and binding of substrates.

D-Ribose 5-phosphate (R5P) allowed us to shed light on the role of position 2 for the binding. R5P proved to be a ligand of Pa-KdsD (Figure 1 E) but not a substrate. The same results were observed for Ec-KdsD.[17] Also in this case, we demonstrated the ability of R5P to compete with the natural substrates at the KdsD active site by competitive STD titration experiments. In fact, with increasing R5P concentration, a decrease in the substrate STD intensities was observed (Figure 2). Finally, Pa-KdsD requires the presence of a phosphate group at position 5 for the substrate recognition. As shown for Ec-KdsD, the nonphosphorylated substrate analogues D-arabinose and

D-ribulose, and also the R5P analogue D-ribose, are not Pa-KdsD ligands (Figure 1 F).

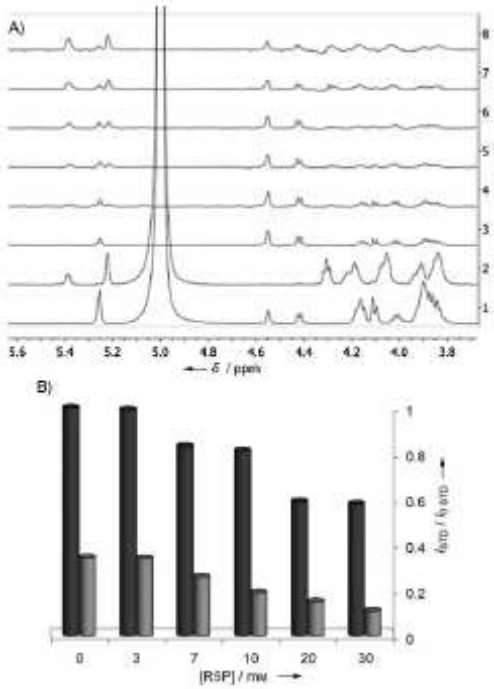


Figure 2. Competitive STD with R5P. A) 1) ^{1}H NMR spectrum of Pa-KdsD/natural substrate mixture at equilibrium; A5P, Ru5P and Pa-KdsD concentrations were 7 mM, 3mM and 30 uM, respectively, NS=16. 2) ^{1}H NMR spectrum of R5P (3.5 mM), NS=16. 3–8) STD spectra of Pa-KdsD/substrate mixture at equilibrium with added R5P at different concentrations: 3) 0 mM, 4) 3mM, 5) 7 mM, 6) 10mM, 7) 20mM, and 8) 30 mM. For all the STD experiments: NS=768, on-resonance frequency= 7.9 ppm, off-resonance frequency= 40 ppm, protein saturation time=2 s. All the samples were dissolved in PBS (pH 7.3, 5°C). B) Decrease in the fractional STD effects of Ru5P H1 (black) and of A5P H1 (grey) as a function of R5P concentration. A5P, Ru5P and Pa-KdsD concentrations were 7 mM, 3mM and 30 uM, respectively. Fractional STD effects were calculated by $(I_0\text{-}I)/I_0$, where $(I_0\text{-}I)$ is the peak intensity in the STD spectrum and I_0 is the peak intensity in the off-resonance spectrum. The largest STD effect ($I_{0\,STD}$) was set to 1 and relative intensities (I_{STD}) were determined.

Mutational analysis of Pa-KdsD—biochemical and NMR characterization

A previous work identified Lys59 and His193 as functionally relevant residues of Ec-KdsD.[16] Moreover, the recently solved crystal structure of

the sugar isomerase domain of the catalytically inactive mutant Ec-KdsD-Lys59Ala allowed the identification of a previously unpredicted His residue (His88, Ec-KdsD numbering), located at the mouth of the active site cavity, as a possible catalytic residue.[36] Ec-KdsD and Pa-KdsD share significant sequence similarity (57% identity and 75% similarity) and Lys59, His88 and His193 are conserved in Pa-KdsD and correspond to Lys56, His85 and His190, respectively. The relevance of the three residues for the enzymatic activity was investigated by converting each one into Ala to abolish its basic or acidic properties. As shown in Table 2, the specific activities of Pa-KdsDLys56Ala, Pa-KdsD-His85Ala and Pa-KdsD-His190Ala were near or below the detection limit of the assay. In view of the drastic losses of activity exhibited by the three mutants, we suggest that all three residues play pivotal roles in KdsD activity. This was further confirmed by *in vivo* complementation studies. The mutant plasmids introduced into the BB-8 conditional mutant strain, in which *kdsD* expression is induced by arabinose and the *gutQ* paralogous copy is inactivated (Experimental Section),[11] were unable to support growth under non-permissive conditions (without arabinose; Table 2). The substrate recognition and binding requirements of Pa- KdsD-Lys56Ala, Pa-KdsD-His85Ala and Pa-KdsD-His190Ala were investigated by STD NMR, as previously described for the wildtype protein. In particular, we verified the interaction of the mutant proteins with the natural substrates A5P and Ru5P and with the substrate analogues D-ribulose, D-arabinose, R5P and Xu5P. From our results, the binding requirements of the three Pa-KdsD mutants are the same as found for the wild-type protein. As a matter of fact, Pa-KdsD-Lys56Ala, Pa-KdsD-His85Ala and Pa-KdsD-His190Ala recognised A5P, Ru5P and R5P but proved unable to bind D-ribulose, D-arabinose and Xu5P (Figure 3).

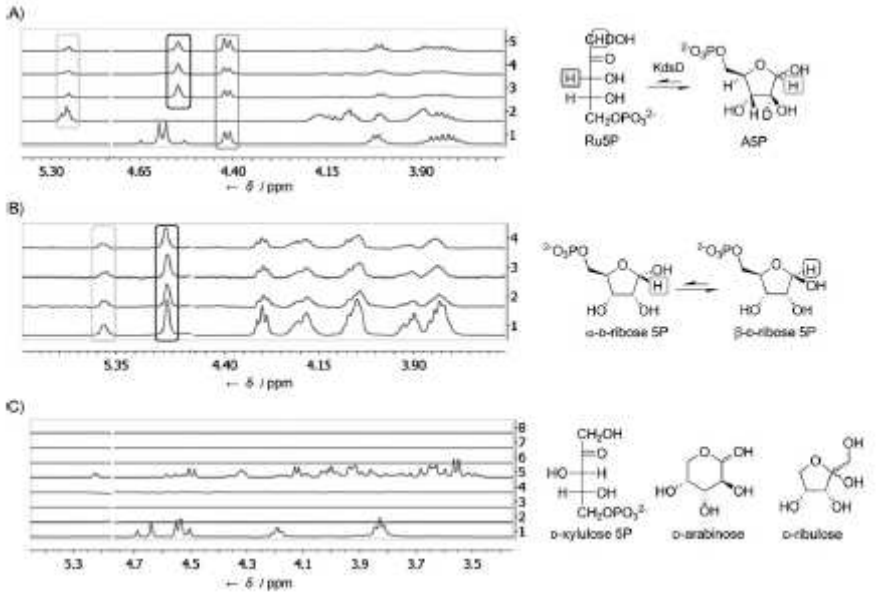


Figure 3. NMR characterisation of the binding requirements of the Pa-KdsD mutants His85Ala, Lys56Ala and His190Ala. A) On the left : ^{1}H NMR spectra 1) of Ru5P and 2) of A5P, and STD spectra 3) of Pa-KdsD-His85Ala/substrate mixture at equilibrium, 4) of Pa-KdsD-Lys56Ala/substrate mixture at equilibrium, and 5) of Pa KdsD-His190Ala/substrate mixture at equilibrium. Each sample contained a different Pa-KdsD mutated protein with a final concentration of 30 uM and substrates at a total concentration of 10 mM. On the right: the reversible isomerisation of Ru5P to A5P catalysed by each Pa-KdsD mutant. When the reaction is carried out in D_2O, deuteration of Ru5P at position 1 and of A5P at position 2 occurs. B) On the left: 1) ^{1}H NMR spectrum of R5P, together with STD spectra 2) of Pa-KdsD-His85Ala/R5P, 3) of Pa-KdsD-Lys56Ala/R5P, and 4) of Pa KdsD-His190Ala/R5P. Each sample contained a different Pa-KdsD mutated protein with a final concentration of 30 uM and R5P (3.5 mM). On the right: R5P exists in solution as a mixture of a and b anomers. C) On the left: 1) ^{1}H NMR spectrum of Xu5P, together with STD spectra 2) of Pa-KdsD-His85Ala/Xu5P, 3) of Pa-KdsD Lys56Ala/Xu5P, and 4) of Pa-KdsD-His190Ala/Xu5P. 5) ^{1}H NMR spectrum of D-arabinose and D-ribulose mixture, together with STD spectra 6) of Pa-KdsD-His85Ala/D-arabinose and d-ribulose mixture, 7) of Pa-KdsD-Lys56Ala/D-arabinose and D-ribulose mixture, and 8) of Pa-KdsD-His190Ala/ D-arabinose and D-ribulose mixture. Each contained a different Pa-KdsD mutated protein with a final concentration of 30 uM and Xu5P (3.5 mM), or a D-arabinose and D-ribulose mixture, in which each compound had a final concentration of 3.5 mM. On the right: structures of Xu5P, D-arabinose and D-ribulose. All the samples were dissolved in PBS (pH 7.3, 5°C). ^{1}H NMR spectra were each recorded with 16 scans. STD spectra were each recorded with 768 scans; protein saturation time=2 s on resonance frequency=7.9 ppm, off-resonance frequency=40 ppm.

At the same time, we verified the mutants' ability to catalyse the isomerisation of the natural substrates A5P and Ru5P. In contrast with the spectrophotometric assay, in which the specific activities of the three proteins were near or below the detection limit, it was possible to study the isomerisation kinetics by ^{1}H NMR. Pa-KdsD mutant specific activities at 378C—determined as increases in the Ru5P H3 integral value versus time per ug of KdsD—were (0.072±0.005) Umg^{-1} for Pa-KdsDLys56Ala, (0.196±0.013) Umg^{-1} for Pa KdsD-His85Ala and (0.043±0.004) Umg^{-1} for Pa-KdsD-His190Ala. These data clearly demonstrate that all the three mutations significantly affect the enzymatic activity, being about 100 times lower for His85Ala, about 200 times lower for Lys56Ala and about 400 times lower for His190Ala, relative to the activity of wild-type Pa-KdsD. We therefore concluded that the amino acids Lys56, His85 and His190 are fundamental for Pa-KdsD activity. The measured residual enzymatic activity could be due to a contribution to catalysis from other residues located very close to the enzymatic active site (assisted catalysis). As a matter of fact, it has been shown that different residues can act as basic and acidic catalytic residues in various sugar isomerases.[19–22] For these reasons, we cannot exclude the possibility that amino acids (e.g., glutamate) located in close proximity to the active site might substitute for the function of the mutated residue, although with low efficiency. The same assay revealed, consistently with what had been observed for the wild-type enzyme, that Pa-KdsD-Lys56Ala, Pa-KdsD-His85Ala and Pa-KdsD-His190Ala are unable to convert R5P into the corresponding ketose. To clarify the roles of the residues Lys56, His85 and His190 further, and in particular to establish whether such amino acids are fundamental for catalysis or for substrate recognition and binding, an additional set of experiments was performed. For this purpose K_D values for the binding of Ru5P to the four enzymes

Pa-KdsD, Pa-KdsD-Lys56Ala, Pa KdsD-His85Ala and Pa-KdsD-His190Ala were determined by the method recently reported by Angulo et al.[37] For each enzyme (protein concentration, 40 uM), a series of STD spectra was recorded in phosphate buffer saline (PBS, pH 7.3), at 5°C with variation of the ligand (A5P/Ru5P 7:3 mixture at equilibrium) concentration. For each Ru5P concentration (0.6, 0.9, 1.2, 1.5, 1.8 mM), STD-AF values for the H3 proton were obtained at different saturation times (0.7, 1.1, 1.5, 2.0, 2.5 s) and fitted by using Equation (1):

$$\text{STD-AF}(t_{sat}) = \text{STD-AF}_{max}\,[1-\exp(-k_{sat}t\text{sat})]$$

The initial slopes—STD-AF_0—were obtained from Equation (2):

$$\text{STD-AF}_0 = \text{STD-AF}_{max}k_{sat}$$

STD-AF_0 values were then represented as a function of the ligand concentration. Subsequent mathematical fitting to a Langmuir isotherm [Eq. (3)]

$$y = B_{max}x/(K_D + x)$$

allowed the B_{max} and K_D values reported in Table 3 to be obtained. Experimental data obtained for Pa-KdsD and their fitting are reported in Figure S3 as an example.

Table 3. Comparison between wild-type Pa-KdsD, Pa-KdsD-Lys56Ala, Pa-KdsD-His85Ala and Pa-KdsD-His190Ala K_D values for the API natural substrate Ru5P by STD NMR.

Protein	B_{max} [a,b]	K_D [mM] [a,b]
Pa-KdsD	4.868 ± 0.133	0.598 ± 0.048
Pa-KdsD-Lys56Ala	4.886 ± 0.182	0.737 ± 0.072
Pa-KdsD-His85Ala	4.591 ± 0.303	1.117 ± 0.156
P-KdsD-His190Ala	5.256 ± 0.227	0.602 ± 0.076

[a] Calculated by the procedure described in ref. [37]. [b] Each measurement is the mean of at least three independent determinations.

Pa-KdsD-Lys56Ala and Pa-KdsD-His190Ala have K_D values of 0.737±0.072 mM and 0.602±0.076 mM, respectively, very similar to the K_D value of 0.598±0.048 mM calculated for Pa-KdsD. Because the K_D value is a good estimation of the enzyme affinity for the substrate Ru5P, our data suggest that both mutant proteins conserve binding affinities comparable to that of wild-type protein. On the other hand, the K_D value for Pa-KdsD-His85Ala is 1.117±0.156 mM, indicating that the corresponding mutation significantly affects the enzyme binding to the natural substrates. Overall our biochemical and NMR data indicate that Lys56, His85 and His190 all play pivotal roles in the activity of Pa-KdsD. Moreover, as a result of the comparison between the K_D values found for the four enzymes, we can infer that the His85 residue is implicated in substrate binding whereas both Lys56 and His190 are not. From these data, we can postulate a structural role for His85 and a catalytic function for Lys56 and His190. Determination of the crystal structures of wild-type and mutant Pa-KdsD would allow confirmation of our speculations.

Conclusions

This work reports the identification and characterisation of *P. aeruginosa* KdsD (Pa-KdsD), a D-arabinose 5-phosphate isomerase involved in the biosynthesis of 3-deoxy-D-*manno*-oct-2-ulosonic acid (Kdo). Kdo is an essential residue of lipopolysaccharide (LPS), a crucial component of the outer membranes of Gram-negative bacteria. We have demonstrated that KdsD is essential for the survival of *P. aeruginosa* and thus represents a key target for the development of novel antibiotic drugs. Pa-KdsD's substrate recognition and binding requirements were characterised by saturation transfer difference (STD) NMR spectroscopy, by study of its interaction with natural substrates and substrate analogues. We found that

Pa-KdsD has substrate recognition and binding properties very similar to those of the *E. coli* homologous enzyme (Ec-KdsD). This is in line with previous data: indeed, Ec-KdsD shows significant sequence similarity to Pa-KdsD and is able to complement a *P. aeruginosa* mutant in which the chromosomal copy of kdsD is disrupted. However, Pa-KdsD proved more demanding with regard to the stereochemistry and to the natures of substituents at position 3 in the substrates. The mutational analysis of Pa-KdsD allowed us to demonstrate that Lys56, His85 and His190 are key residues for enzyme activity, because their conversion into Ala causes 100- to 400-fold decreases in mutant Pa-KdsD specific activity. In addition, STD experiments to compare the affinities of wild-type and mutant enzymes for the natural substrates were set up. The significant differences between enzyme affinities led us to postulate a structural role for His85 and a catalytic function for Lys56 and His190.

The biochemical and NMR structural data reported here not only shed light on the function and enzymatic activity of *P. aeruginosa* KdsD, but also provide important information for the rational design of new KdsD inhibitors. Similarities and differences between Pa-KdsD and Ec-KdsD should be exploitable to design both novel specific or broad-spectrum antibiotics against pathogenic strains of *P. aeruginosa* and *E. coli*.

Experimental Section

Bacterial strains and plasmids: The bacterial strains and plasmids used in this study are listed in Table S1. *E. coli* and *P. aeruginosa* strains were grown in LD medium.[38] Cultures were grown at 37°C with vigorous aeration. Oligonucleotides are listed in Table S2.

Cloning, over-expression and purification of Pa-KdsD: The Pa-*kdsD* gene was cloned according to the deposited reading frame of the annotated

genomic sequence to encode a protein of 326 amino acids (http://www.pseudomonas.com). The recombinant Pa-KdsD protein was over-expressed from a T5 promoter (pQE31, Table S1), which allows the expression of the protein with an N-terminal His tag by use of the oligonucleotides primers PAO fw BamHI and PAO rev PstI (Table S2). The expression construct was initially transferred into *E. coli* XL-1Blue cells for plasmid amplification and then to *E. coli* M15/pREP4 cells for protein expression. Wild-type Pa-KdsD and Pa-KdsD mutants were purified by Ni-chelate affinity chromatography. *E. coli* cells were grown at 37°C with shaking in LD medium containing ampicillin (100 ug mL^{-1}) and kanamycin (25 ug mL^{-1}), until the OD_{600} reached 0.6. Induction was then carried out for 4 h with IPTG (0.5 mM). Cells were harvested, washed with sodium phosphate (20 mM, pH 7.0) and resuspended in three volumes of sodium phosphate (50 mM, pH 8.0), NaCl (300 mM) and phenylmethylsulfonyl fluoride (1 mM). After treatment with lysozyme (0.3 ug mL^{-1}) for 30 min, cells were disrupted by sonication and centrifuged at 40000*g* for 30 min. The supernatant was loaded at a flow rate of 0.5 $mLmin^{-1}$ onto a Ni-NTA Agarose (Qiagen) column (bed volume 5 mL) pre-equilibrated with sodium phosphate (50 mM, pH 8.0) and NaCl (300 mM). The column was washed with 10 volumes of sodium phosphate (50 mM, pH 8.0), NaCl (300 mM) and imidazole (10 mM) and the enzyme was eluted with a 50 mL stepwise imidazole gradient (100 to 300 mM) in the same buffer. Fractions (5 mL) were collected and those displaying the highest purity (>95%, as assessed by SDS-PAGE analysis) were dialysed in sodium phosphate (50 mM, pH 8.0) and NaCl (300 mM). Protein aliquots (600 ug each) were lyophilised and stored at -80°C.

Site-directed mutagenesis of Pa-*kdsD* and complementation assay: Site-directed mutagenesis in Pa-*kdsD* was performed by PCR with use of the

QuickChange site-directed mutagenesis kit (Stratagene) and the oligonucleotides listed in Table S1, with the pET14 Pa-*kdsD* plasmid as a template. The plasmids were amplified in *E. coli* XL1Blue and then introduced by transformation in the BB-8 mutant strain.[11] Complementation analysis was performed as described elsewhere,[16] with the exception that the bacterial suspensions were replicated on LD agar plates containing ampicillin (100 mg mL^{-1}) and either supplemented or not supplemented with arabinose (0.2 %).

Determination of protein activity: Wild-type and mutant Pa-KdsD activities were assayed continuously with use of A5P (Sigma–Aldrich) as a substrate. The assay mixture contained Tris·HCl, (pH 8.5, 50 mM) and A5P (5 mM). After enzyme addition the increase in absorbance was monitored at 37°C at 280 nm.[39] At this wavelength, the molar absorption coefficient for Ru5P was 58.6 $M^{-1}cm^{-1}$. One unit of enzyme is defined as the amount of enzyme that produces 1 umol of product per min under the standard assay conditions.

Conjugation experiments: PAOΔAPI was constructed by use of the suicide vector pDM4, which carries the chloramphenicol resistance gene, and the *sacB* gene, which confers sucrose sensitivity. A 250 bp sequence internal to the sequence of *kdsD* gene coding for the SIS domain was amplified with use of the oligonucleotides PAO 250 fw KpnI and PAO 250 rev SacI (Table S2) and cloned into the pSD-1 vector downstream of the araBp promoter to create the plasmid pSD-1 Pa-*kdsD*. pDM4 Pa-*kdsD* plasmid was created by amplifying the cassette *araC-araBp* Pa-*kdsD* from pSD-1 Pa-*kdsD* with use of the oligonucleotides *araC* fw XbaI and PAO 250 rev *SacI* (Table S2) and cloning into pDM4. pDM4 Pa-*algR* was created by the same strategy (Table S1). Plasmid pDM4 Pa-kdsD and pDM4 Pa-*algR* were conjugated from *E. coli* strain S17–1 (*λpir*) into PAO1 as described below.

Briefly, PAO1/Ec *kdsD* and *E. coli* S17–1 *λpir*/pDM4 *kdsD* or S17–1 *λpir*/pDM4 *algR* were grown overnight in LD medium containing chloramphenicol (25 ug mL^{-1}) at 37°C. A PAO1/Ec-*kdsD* culture (50 mL) grown to an OD_{600} of 4 was mixed with *E. coli* S17–1 *λpir*/pDM4 *kdsD* or S17–1 *λpir*/pDM4 *algR* cultures (50 mL) grown at the same OD_{600} in $MgSO_4$ (10 mM, 5 mL). The suspension was filtered (membrane filter Millipore 0.45 um HA) to concentrate the cells and each filter was deposed onto LD agar plates supplemented with arabinose (0.2 %). The plates were incubated overnight at 37°C and the bacteria were then resuspended in $MgSO_4$ (10 mM, 3 mL) by submerging the filters in the solution and shaking vigorously. For each conjugation, four 100-fold dilutions (100 uL) in a final volume of 1 mL were plated onto *Pseudomonas* isolation agar plates supplemented with carbenicillin (200 mg mL^{-1}), chloramphenicol (85 $ugmL^{-1}$) and arabinose (0.2%) and incubated overnight at 37°C. Mutants arising through single crossover events were then screened by PCR.

NMR spectroscopy experiments: All of the experiments were recorded with a Varian Mercury (400 MHz) or a Bruker Advance III (600 MHz) instrument with a cryoprobe. The ligand resonances were assigned by $^{1}H,^{1}H$ COSY and $^{1}H,^{13}C$ HSQC NMR spectroscopy.

A basic 1D-STD sequence was used with the on-resonance frequency of 7.9 ppm and the off-resonance frequency of 40 ppm. A train of Gaussian-shaped pulses of 50 ms each was employed, with total saturation times of the protein envelope of 5, 3, 2, 1.5, 1 or 0.5 s. The total saturation time was adjusted by the number of shaped pulses. A T1ρ filter of 2 ms was employed to eliminate the background signals from the protein. All of the samples were dissolved in deuterated PBS (pH 7.3) at 5°C. Total sample volumes were 550 uL. The on- and off-resonance spectra were acquired simultaneously with the same number of scans. The STD spectrum was

obtained by subtraction of the on-resonance spectrum from the off-resonance spectrum. Subtraction was performed by phase cycling to minimise artefacts arising from magnet and temperature instabilities.[40] Reference experiments with samples containing only the free compounds tested were performed under the same experimental conditions to verify true ligand binding. The effects observed in the presence of the protein were due to true saturation transfer, because no signals were present in the STD spectra obtained in the reference experiments, except for residues from HDO, indicating that artefacts from the subtraction of compound signals were negligible.

API specific activities were determined at 37°C as increases in the H3 (Ru5P and 3-deoxy-Ru5P) or H4 (4-deoxy-Ru5P) integral values versus time per mg of Pa-KdsD. Spectra were collected over time (673 s for wild-type Pa-KdsD with A5P and 4-deoxy-A5P, 937 s for wild-type Pa-KdsD with 3-deoxy-A5P, 560 s for Pa-KdsD-His85Ala with A5P, 1900 s for Pa-KdsD-His190Ala with A5P and 1300 s for Pa-KdsD-Lys56Ala) at intervals of 33 s. Each measurement is the mean of at least three independent determinations. The integral values were then converted into Ru5P concentrations and enzymatic activities calculated. One unit of enzyme is defined as the amount of enzyme that produces 1 umol of product per min under the standard assay conditions.

Spectra processing and analysis were performed with the software MestReNova (http://mestrelab.com/). The fractional STD effect was calculated by use of $(I_0\text{-}I)/I_0$, in which $(I_0\text{-}I)$ is the peak intensity in the STD spectrum and I_0 is the peak intensity in the off-resonance spectrum.[30]

For titration experiments with R5P, the largest STD effect ($I_{0\ STD}$) in the STD spectrum of KdsD and natural substrate mixture was set to one and the relative intensities (I_{STD}) for the other signals were determined.

K_D values for the binding of Ru5P to the four enzymes Pa-KdsD, Pa-KdsD-Lys56Ala, Pa-KdsD-His85Ala or Pa-KdsD-His190Ala were determined by application of the method recently reported by Angulo et al.[37] Different concentrations of a solution of A5P/Ru5P 7:3 were added at 5°C to an enzyme solution (40 uM in PBS, pH 7.3). For each concentration, the mixture was allowed to reach equilibrium for 15 min and then a series of STD spectra was recorded, with variation of the saturation time. Both ligand concentration and saturation time parameters were optimised. For each Ru5P concentration (0.6, 0.9, 1.2, 1.5 or 1.8 mM), STD-AF values for the H3 proton were obtained at different saturation times (0.7, 1.1, 1.5, 2.0 and 2.5 s) and fitted by use of Equation (1). The initial slopes— STD-AF_0—were obtained from Equation (2). STD-AF_0 values were then represented as a function of the ligand concentration, and mathematical fitting to a Langmuir isotherm [Eq. (3)] allowed the B_{max} and K_D values to be obtained.

All of the compounds tested as potential KdsD ligands were purchased from Sigma–Aldrich, with the exceptions of d-ribulose, which was purchased from Carbosynth, and of 3-deoxy-A5P and 4-deoxy-A5P, which were synthesised by us.

Acknowledgements

This work was supported by the Italian Cystic Fibrosis Research Foundation (Grant FFC#10/2008) with contributions from Ferrara and Bologna FFC Delegations, Comacchio Supporting Group, and Associazione Trentina FC. S.S. is a recipient of a fellowship from the Italian Cystic Fibrosis Research Foundation, and PRIN2008, “Antagonist and Inhibitors of LPS Biosynthesis: Design, Synthesis and Biological Validation”.

Keywords: antibiotics · enzymes · inhibitors · isomerization · NMR binding studies

[1] T. Ikeno, K. Fukuda, M. Ogawa, M. Honda, T. Tanabe, H. Taniguchi, Microbiol. Immunol. 2007, 51, 929 –938.
[2] F. Yoshimura, H. Nikaido, J. Bacteriol. 1982, 152, 636– 642.

[3] K. Poole, R. Srikumar, Curr. Top. Med. Chem. 2001, 1, 59–71.
[4] G. B. Pier, Int. J. Med. Microbiol. 2007, 297, 277– 295.
[5] M. Y. Lam, E. J. McGroarty, A. M Kropinski, L. A. MacDonald, S. S. Pedersen, N. Høiby, J. S. Lam, J. Clin. Microbiol. 1989, 27, 962 –967.
[6] M. Rivera, L. E. Bryan, R. E. Hancock, E. J. McGroarty, J. Bacteriol. 1988, 170, 512–521.
[7] H. L. Rocchetta, L. L. Burrows, J. S. Lam, Microbiol. Mol. Biol. Rev. 1999, 63, 523–553.
[8] J. D. King, D. Koc_ncov_, E. L. Westman, J. S. Lam, Innate Immun. 2009, 15, 261–312.
[9] T. C. Meredith, R. W. Woodard, J. Biol. Chem. 2003, 278, 32771– 32777.
[10] T. C. Meredith, R. W. Woodard, J. Bacteriol. 2005, 187, 6936– 6942.
[11] P. Sperandeo, C. Pozzi, G. Deh_, A. Polissi, Res. Microbiol. 2006, 157, 547 –558.
[12] C. R. Raetz, C. Whitfield, Annu. Rev. Biochem. 2002, 71, 635 –700.
[13] C. A. Schnaitman, J. D. Klena, Microbiol. Rev. 1993, 57, 655– 682.
[14] A. G. Walsh, M. J. Matewish, L. L. Burrows, M. A. Monteiro, M. B. Perry, J. S. Lam, Mol. Microbiol. 2000, 35, 718 –727.
[15] R. Rahim, L. L. Burrows, M. A. Monteiro, M. B. Perry, J. S. Lam, Microbiology 2000, 146, 2803 – 2814.
[16] S. Sommaruga, L. De Gioia, P. Tortora, A. Polissi, Biochem. Biophys. Res. Commun. 2009, 388, 222 –227.
[17] C. Airoldi, S. Sommaruga, S. Merlo, P. Sperandeo, L. Cipolla, A. Polissi, F. Nicotra, Chem. Eur. J. 2010, 16, 1897– 1902.
[18] I. A. Rose, Adv. Enzymol. Relat. Areas Mol. Biol. 1975, 43, 491– 517.
[19] B. Asb_th, G. N_ray-Szab_, Curr. Protein Pept. Sci. 2000, 1, 237– 254.
[20] C. J. Jeffery, R. Hardr_, L. Salmon, Biochemistry 2001, 40, 1560– 1566.
[21] J. H. Lee, K. Z. Chang, V. Patel, C. J. Jeffery, Biochemistry 2010, 49, 7799 –7805.
[22] A. Teplyakov, G. Obmolova, M. A. Badet-Denisot, B. Badet, I. Polikarpov, Structure 1998, 6, 1047 –1055.
[23] P. L. Taylor, K. M. Blakely, G. P. de Leon, J. R. Walker, F. McArthur, E. Evdokimova, K. Zhang, M. A. Valvano, G. D. Wright, M. S. Junop, J. Biol. Chem. 2008, 28, 32835– 32845.
[24] N. J. Harmer, J. Mol. Biol. 2010, 400, 379– 392.
[25] P. Sperandeo, G. Dehç, A. Polissi, Biochim. Biophys. Acta Mol. Cell Biol. Lipids 2009, 1791, 594– 602.
[26] L. A. Morici, A. J. Carterson, V. E. Wagner, A. Frisk, J. R. Schurr, K. Hçner zu Bentrup, D. J. Hassett, B. H. Iglewski, K. Sauer, M. J. Schurr, J. Bacteriol. 2007, 189, 7752 – 7764.
[27] M. Mayer, B. Meyer, Angew. Chem. 1999, 111, 1902 – 1906; Angew. Chem. Int. Ed. 1999, 38, 1784 –1788.
[28] B. Meyer, T. Peters, Angew. Chem. 2003, 115, 890 –918; Angew. Chem. Int. Ed. 2003, 42, 864– 890.
[29] T. Biet, T. Peters, Angew. Chem. 2001, 113, 4320– 4323; Angew. Chem. Int. Ed. 2001, 40, 4189 – 4192.
[30] O. Berteau, C. Sandstrçm, J. Bielicki, D. S. Anson, L. Kenne, J. Am. Chem. Soc. 2003, 125, 15296– 15297.

[31] Y. Yuan, X. Wen, D. A. R. Sanders, B. M. Pinto, Biochemistry 2005, 44, 14080 – 14089.
[32] L. Brecker, G. D. Straganz, C. E. Tyl, W. Steiner, B. Nidetzky, J. Mol. Catal. B 2006, 42, 85– 89.
[33] Y. Yuan, X. Wen, D. A. R. Sanders, B. M. Pinto, J. Am. Chem. Soc. 2008, 130, 3157– 3168.
[34] C. Airoldi, S. Merlo, F. Nicotra, J. Carbohydr. Chem. 2010, 29, 30– 38.
[35] A. Kohen, A. Jackob, T. Baasov, Eur. J. Biochem. 1992, 208, 443– 449.
[36] L. J. Gourlay, S. Sommaruga, M. Nardini, P. Sperandeo, G. Deh_, A. Polissi, M. Bolognesi, Protein Sci. 2010, 19, 2430 –2439.
[37] J. Angulo, P. M. Enriquez-Navas, P. M. Nieto, Chem. Eur. J. 2010, 16, 7803 –7812.
[38] D. Ghisotti, R. Chiaramonte, F. Forti, S. Zangrossi, G. Sironi, G. Dehò, Mol. Microbiol. 1992, 6, 3405 –3413.
[39] F. C. Knowles, M. K. Pon, N. G. Pon, Anal. Biochem. 1969, 29, 40– 47.
[40] M. Mayer, B. Meyer, J. Am. Chem. Soc. 2001, 123, 6108 – 6117.

Supporting Information

Organization of the *yrbG-lptB* cluster in *E. coli* K12 and comparison with the homologous region in *P. aeruginosa* PAO1.

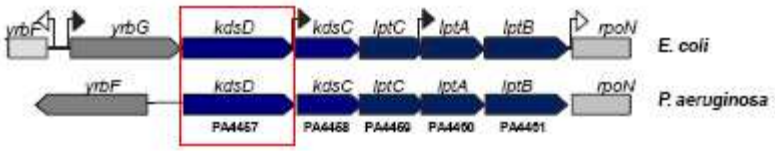


Figure S1. The annotated ORFs are indicated below each gene.

STD-NMR on Pa-KdsD and natural substrate mixture with STD build-up curves.

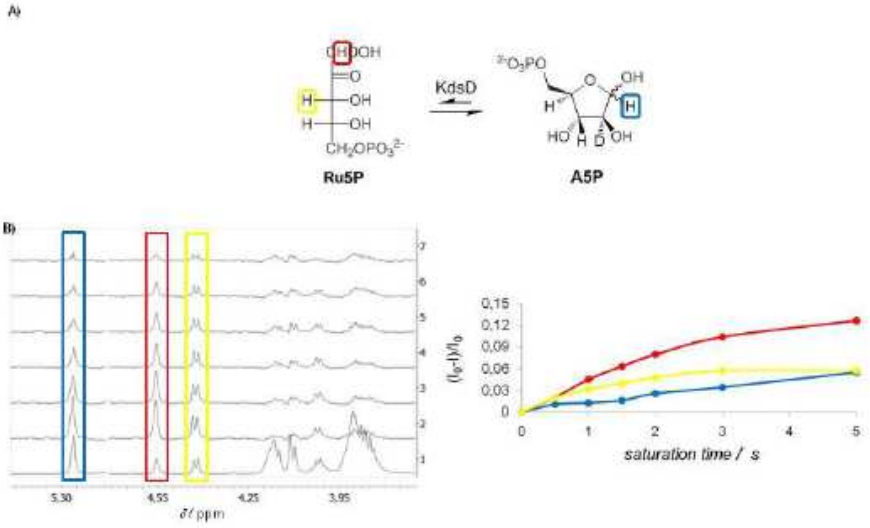


Figure S2. A) The reversible isomerization of Ru5P to A5P catalyzed by KdsD. B) On the left: ^{1}H-NMR spectrum of Pa-KdsD-substrate mixture at equilibrium (A5P:Ru5P 7:3), number of scans (NS)=16 (1); STD spectra of KdsD-substrate mixture at equilibrium, at a Pa-KdsD:A5P 1:233 molar ratio and a Pa-KdsD:Ru5P 1:100 molar ratio, recorded with different protein saturation times: 5 s (2); 3 s (3); 2 s (4); 1.5 s (5); 1s (6); 0.5 s (7), NS=768, on-resonance frequency=7.9 ppm, offresonance frequency=40 ppm. Spectra 1-7 were recorded on the same sample; sample was dissolved in PBS, pH 7.3, 5°C. On the right: Fractional STD effects for A5P H1 (blue), Ru5P H1 (red) and Ru5P H3 (yellow) calculated by $(I_0\text{-}I)/I_0$, where

(I_0-I) is the peak intensity in the STD spectrum and I0 is the peak intensity in the off resonance spectrum.

Table S1. Bacterial strains and plasmids.

Strain or plasmid	Relevant characters*/origin	Source or reference
E. coli		
XL-1 Blue	*recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac*	(Stratagene)
BB-8	BW25113 ⊕(*kan-araC-araBp*kdsD)1 Δ*gutQ*1::*cat*	(Sperandeo *et al.*, 2006)
M15/pREP4	F$^-$ *lac thi mtl*/pREP4	(Qiagen)
BW25113	*lacI^q rrnBT14 lacZWJ16 hsdR514 araBADAH33 rhaBADLD78*	(Datsenko and Wanner, 2000)
S17-1 (λ*pir*)	*thi pro hsdR$^-$M$^+$ recA* RP4-2::Tc::Mu-Km::Tn*7*, lysogenic for λ*pir*	(Simon *et al.*, 1983)
P. aeruginosa		
PAO1	*Pseudomonas aeruginosa* wild type	(Stover *et al.*, 2000)
PAOΔAPI	PAO1 *kdsD* (*araC-araBp*Pa-*kdsD*)1/ pVI533 Ec-*kdsD*	This study
Plasmids		
pET14b	Ap^R,ColEI *ori*, T7 promoter	(Novagen)
pET14 Pa-*kdsD*	pET-14 derivative, overexpresses *E. coli* KdsD upon IPTG induction	This study
pET14 Pa*kdsD* His56Ala	pET14 Pa-*kdsD* derivative, carries H56A aminoacid substitution	This study
pET14 Pa*kdsD* His85Ala	pET14 Pa-*kdsD* derivative, carries H85A aminoacid substitution	This study
pET14 Pa*kdsD* His190Ala	pET14 Pa-*kdsD* derivative, carries H190A aminoacid substitution	This study
pQE-31	Ap^R, ColEI *ori*, T5 promoter.	(Qiagen)
pQE-31 Pa *kdsD*	pQE-31; overexpresses *E. coli* KdsD upon IPTG induction	This study
pSD-1	Carries the TIR downstream of *araBp*. Source of *kan-araC-araBp* cassette for gene fusion.	(Sperandeo *et al.*, 2006)
pSD-1 Pa-*kdsD*	pSD1 derivative, carries 250 bp of Pa-*kdsD* downstream *araBp*	This study
pSD-1 Pa-*algR*	pSD1 derivative, carries 250 bp of Pa-*algR* downstream *araBp*	This study
pDM4	Cm^R, sacB; allelic exchange vector	(Milton *et al.* 1996)
pDM4 Pa-*kdsD*	pDM4 containing *araC-araBp* cassette upstream of 250 bp of Pa-*kdsD*	This study
pDM4 Pa-*algR*	pDM4 carrying *araC-araBp* cassette upstream of 250 bp of Pa-*algR*	This study
pVI533	Cb^R, pMMB66EH with *araC/araBp* promoter	(Sze and Shingler, 1999)
pVI533 Ec-*kdsD*	pVI533 derivative, carries Ec-*kdsD* under the control of *araBp* promoter	This study

*Cb, carbenicillin; Cm, chloramphenicol; Km, kanamycin; Ap, ampicillin.

References

- Datsenko, K.A., Wanner, B.L. (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products, Proc. Natl. Acad. Sci. USA 97, 6640–6645.
- Milton, D. L., R. O'Toole, P. Horstedt, and H. Wolf-Watz. 1996. Flagellin A is essential for the virulence of *Vibrio anguillarum*. J. Bacteriol. 178:1310-1319.

- Simon, R., Priefer, U., Püler, A. (1983) A broad host range system for *in vivo* genetic engineering: Transposon mutagenesis in Gram-negative bacteria, Biotechnology 1, 417–435.
- Sperandeo, P., Pozzi, C., Dehò, G., and Polissi, A. (2006) Non-essential KDO biosynthesis and new essential cell envelope biogenesis genes in the *Escherichia coli yrbG-yhbG* locus *Res.Microbiol.* 157, 547-558.
- Stover, C. K., X. Q. Pham, A. L. Erwin, S. D. Mizoguchi, P. Warrener, M. J. Hickey, F.S. L. Brinkman, W. O. Hufnagle, D. J. Kowalik, M. Lagrou, R. L. Garber, L. Goltry, E. Tolentino, S. Westbrock-Wadman, Y. Yuan, L. L. Brody, S. N. Coulter, K. R. Folger, A. Kas, K. Larbig, R. Lim, K. Smith, D. Spencer, G. K.-S. Wong, Z. Wu, I. T. Paulsen, J. Reizer, M. H. Saier, Hancock, R. E. W., Lory S., Olson, M. V. (2000). Complete genome sequence of *Pseudomonas aeruginosa* PAO1, an opportunistic pathogen. Nature 406, 959–964.
- Sze, C.C., Shingler, V. (1999) The alarmone (p)ppGpp mediates physiological-responsive control at the sigma 54-dependent Po promoter. Mol Microbiol 31,1217–1228.

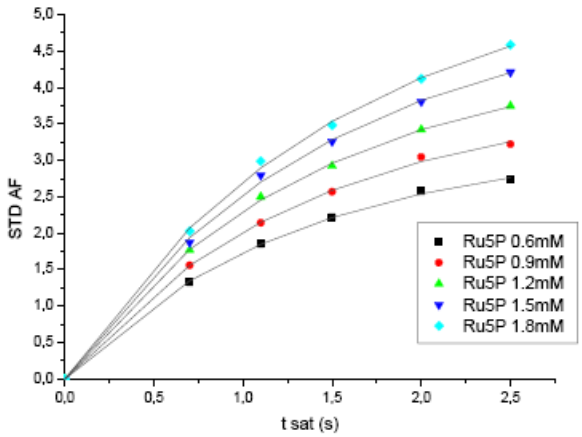
Table S2. Oligonucleotides.

Name	Sequence[a] (5' to 3')	Use/description
PAO fw *Bam*HI	gaagagGGATCCGAACATGAGCCAGAACCTCG	pQE-31 Pa *kdsD* with PAO PAO rev *Pst*I
PAO rev *Pst*I	gaagatCTGCAGTTACATCACTCCTGCGCGC	pQE-31 Pa *kdsD* with PAO PAO fw *Bam*HI
PAO fw *Nde*I	gaagatCATATGAACATGAGCCAGAACCTCG	pET14 Pa-*kdsD* with PAO rev *Bam*HI
PAO rev *Bam*HI	gaagttGGATCCTTACATCACTCCTGCGCGC	pET14 Pa-*kdsD* with PAO rev *Nde*I
Lys56Ala fw	GTCGTCGGCATGGGCGCATCCGGCCACGTCG	pET14 Pa*kdsD* Lys56A with Lys56Ala rev
Lys56Ala rev	CGACGTGGCCGGATGCGCCCATGCCGACGAC	pET14 Pa*kdsD* His56A with Lys56Ala fw
His85Ala fw	CCGGCCGAGGCCAGCGCTGGCGACATGGGCATG	pET14 Pa*kdsD* His56A with H85Ala rev
His85Ala rev	CATGCCCATGTCGCCAGCGCTGGCCTCGGCCGG	pET14 Pa*kdsD* His56A with H85Ala fw
His190Ala fw	GACTTCGCCTTCTCGGCCCCGGGCGGTGCGCTG	pET14 Pa*kdsD* His190A with H190Ala rev
His190Ala rev	CAGCGCACCGCCCGGGGCCGAGAAGGCGAAGTC	pET14 Pa*kdsD* His56A with H190Ala fw
PAO 250 fw *Kpn*I	gagaatGGTACCCACGTCGGCAAAAAGATCGC	pSD-1 Pa-*kdsD* with PAO 250 rev *Sac*I
PAO 250 rev *Sac*I	gaatatGAGCTCGCGTCCAGGTTCACTTCCGC	pSD-1 Pa-*kdsD* with PAO 250 fw *Kpn*I
algR 250 fw *Xba*I	gaggatTCTAGACGGCCAGACTCTGCGAGCGG	pSD-1 Pa-*algR* with algR 250 rev *Sac*I
algR 250 rev *Sac*I	accactGAGCTCCTCGATCCCCTTGCGGGTCC	algR 250 rev *Xba*I with algR 250 *Xba*I
araC fw *Xba*I	gaggatTCTAGAGCCGTCAATTGTCTGATTCG	*araC-araBp kds*D 250 fragment with PAO 250 rev *Sac*I; *araC-araBp algR* 250 with algR 250 rev *Sac*I

[a] Upper case, *P. aruginosa* genomic sequences; lower case, not homologous sequences; underlined upper case, restriction sites or codons changed by site-directed mutagenesis.

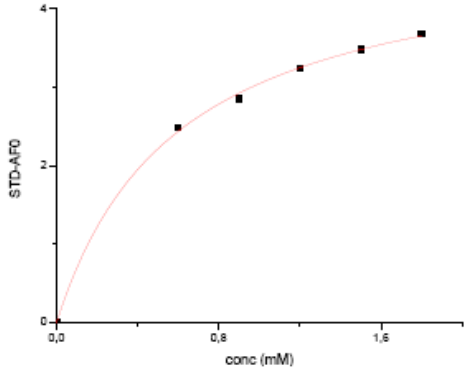
Calculation of K_D value for Ru5P binding to Pa-KdsD by the binding isotherm of STD-AF initial growth rates approach.

A)



[Ru5P] (mM)	STD-AF$_{max}$	k_{sat}	Reduced Chi-Sqr	Adj. R-Square
0,6	3,23722±0,06224	0,76923±0,03039	6,68E-04	0,99935
0,9	3,87178±0,09566	0,73713±0,03645	0,00138	0,99903
1,2	4,47137±0,08057	0,72495±0,02588	9,27E-04	0,9995
1,5	5,16441±0,17848	0,67404±0,04421	0,00359	0,99848
1,8	5,71324±0,18574	0,64389±0,03872	0,00334	0,9988

B)



	B_{max}	K_D	Reduced Chi-Sqr	Adj. R-Square
$STD\text{-}AF_0$	4,86804±0,13327	0,59844±0,04822	0,00209	0,99886

Figure S3. K_D value calculation by the binding isotherm of STD-AF initial growth rates approach. A) For each Ru5P concentration (0.6, 0.9, 1.2, 1.5, 1.8 mM; enzyme concentration 40 µM), STD-AF values for H3 proton were obtained at different saturation times (0.7, 1.1, 1.5, 2.0, 2.5 s) and fit by using the equation $STDAF(t_{sat})=STD\text{-}AF_{max}[1-\exp(-k_{sat}t_{sat})]$. B) The initial slopes, $STD\text{-}AF_0$, were obtained from $STD\text{-}AF_0=STD\text{-}AF_{max}k_{sat}$. b) $STD\text{-}AF_0$ values were then represented as a function of the ligand concentration, and the mathematical fit to a Langmuir isotherm ($y=B_{max}x/(K_D+x)$) allowed to obtain B_{max} and K_D values.

Going towards the development of novel inhibitors of *Pseudomonas aeruginosa* D-Arabinose-5P Isomerase

Introduction

Through the NMR interaction studies we performed on *E. coli* and *P. aeruginosa* KdsD and a small panel of substrate analogues and synthetic compounds,[1,2] we provided very important information about substrate binding mode:

1. the correct stereochemistry at position 2 is not needed for substrate recognition, but it is needed for the enzymatic catalysis, as H2 must be extracted by a suitably positioned basic residue in the enzymatic pocket;
2. position 3 is involved in recognition and therefore the correct stereochemistry is fundamental for enzyme binding;
3. the hydroxyl group at position 4 is needed neither for binding nor for substrate conversion, as demonstrated with 4-deoxyA5P, that is converted to the corresponding ketose in the presence of KdsD.
4. the phosphate group is a fundamental requirement for enzyme recognition and binding.

In order to better characterise KdsD and get further insights of the enzymatic binding requirements needed for the design of more potent inhibitors, we undertook a molecular recognition study of two of the most potent inhibitors found to date (compounds **1** and **2**, Fig. 1 B)[3] together with some other arabinose derivatives (compounds **3**-**9**, Fig.1 B).

The ability of compound **1**-**9** to recognize and bind API was investigated by Saturation Transfer Difference (STD)-NMR spectroscopy, according to the same procedure previously described [1,2]; furthermore the newly

synthesised arabinose derivatives **8-9** were also characterised for their activity *in vivo*.

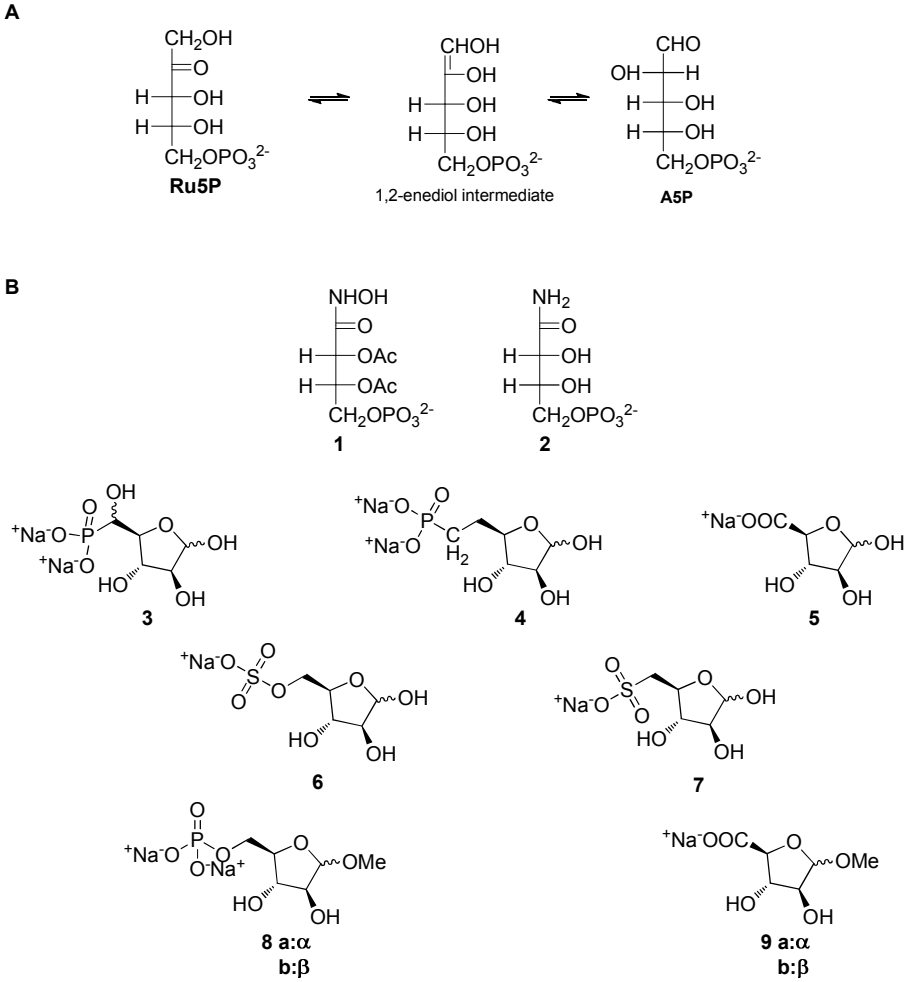


Figure 1. (**A**) proposed KdsD isomerisation mechanism; (**B**) structures of the compounds tested in this study.

Results and discussion

Carbohydrate analogues have been chosen as potential inhibitors, since we hope to take advantage of the usual uptake channels of the bacteria to deliver the drug into the cell. Compounds **1** and **2** were designed on the basis of the presumed high-energy 1,2-enediol intermediate involved in the

enzymatic inter-conversion of Ru5P and A5P (Fig.1 A); compounds **8**-**9**, that are blocked into their cyclic forms as methyl arabinosides, were synthesised and studied in order to shed light on the possibility that KdsD might accommodate also the substrate in its hemiacetalic form; moreover mimetics obtained by substitution of the phosphate group with other non-hydrolizable phosphate derivatives and acidic moieties were taken into consideration in order to understand if a these groups might be good and/or stable mimetics of the natural phosphate counterpart.

The preliminary ^{1}H-NMR analysis of a solution of inhibitor **7** in PBS, pH 7.4, 25°C, revealed the presence of a mixture of compounds, interconverting over time (Figure 2).

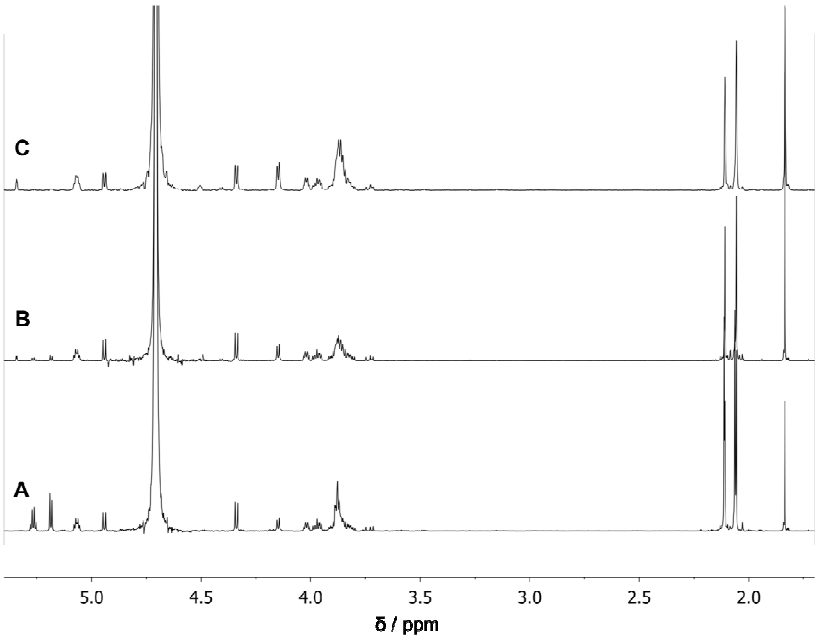


Figure 2. ^{1}H spectra recorded on a mixture containing compound **7** dissolved in PBS, pH 7.0, 25° C recorded immediately after dissolution (**A**), after 1 day (**B**) or after 2 days (**C**).

The careful analysis of $^{1}H,^{1}H$-COSY, $^{1}H,^{1}H$-TOCSY and $^{1}H,^{13}C$-HSQC spectra of the mixture, acquired at different times, allowed the identification of the diverse components, differing for the number and position of acetyl groups, whose structure are reported in Figure 3.

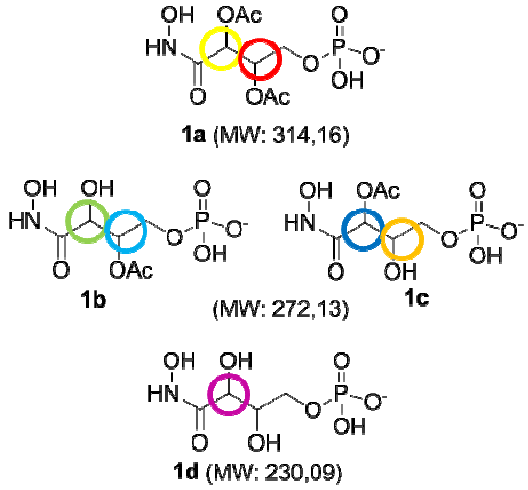


Figure 3. Structure of the different species found in a solution of compound 1 dissolved in PBS, pH 7.0, 25°C.

NMR data were supported by mass spectroscopy analysis, that found the masses 314.1, 272.0 and 230.0, corresponding to the di-acetylated (**1a**), the two mono-acetylated (**1b** and **1c**) and the de-acetylated (**1d**) products respectively. Thus, compound **1** undergoes spontaneous deacetylation in aqueous solution, generating the corresponding mono and de-acetylated species. STD-NMR experiments allowed to identify which mixture components were able to bind KdsD. The STD spectra were acquired on samples dissolved in PBS, pH 7.0, 25°C and prepared by adding the enzyme to different aliquots of a solution of compound **1** (5 mM) after three incubation periods (0, 2 and 3 days). The relative ratio of the species derived from molecule **1** was different for each sample and allowed to

screen all the compounds present in the mixture as potential KdsD ligands. Figure 4 shows ^{1}H and STD spectra recorded for each incubation period.

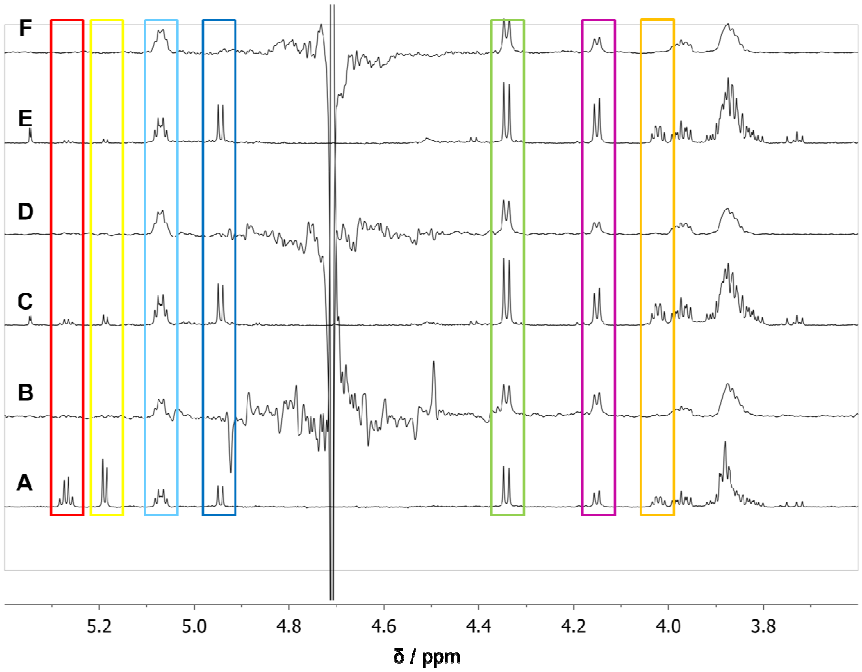


Figure 4. **A,C,E**) ^{1}H NMR spectra of a mixture containing compound 7 (5 mM) in PBS, pH 7.0, 25°C immediately its dissolution (**A**), after 2 days (**C**) and after 3days (**E**), to which 30 µM KdsD was added. B,D,F) STD NMR spectra recorded on the same samples immediately after compound 7 dissolution (**B**), after 2 days (**D**) and after 3days (**F**) with a saturation time of 3 s and a saturation frequency of 7.9 ppm.

These experiments clearly indicated that only molecules **1b** and **1d** interact with the enzyme. In fact, some resonances belonging to these compounds appeared in the STD spectra, while compound **1a** and **1c** signals are absent. At variance with **1b** and **1d**, both **1a** and **1c** are acetylated in position 3, suggesting that this modification is not compatible with the interaction with KdsD. These results are in whole agreement with our previous findings, according to which position 3 is fundamental for enzyme recognition and binding. In the light of these data, we could deduced that molecule **1a**, previously described as KdsD inhibitor [3], is not the active specie; instead,

the active compounds are probably represented by molecules **1b** and **1d**, thus owning a potency higher than that formerly reported, as the actual tested concentrations were lower than expected.

The interaction with KdsD of compound **2** was verified in a similar manner. ^{1}H NMR analysis revealed the presence of a second species (Figure 5A). Unfortunately, its nature was not identified, due to its/their very low solubility (very broad resonances were obtained for concentrations higher than 2 mM, see Figure 5B) and the inability to measure molecular weights by mass spectrometry.

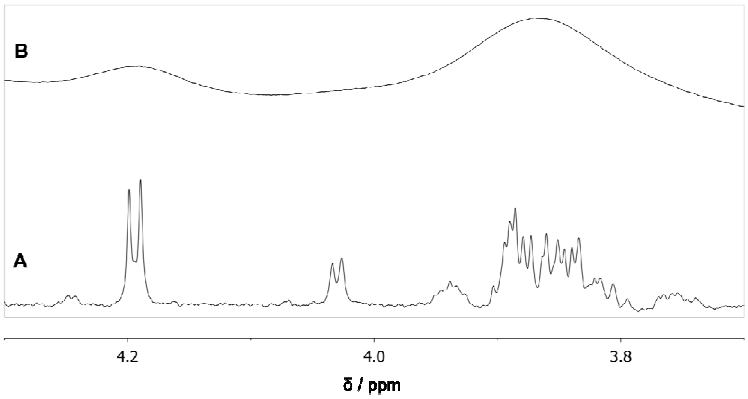


Figure 4. ^{1}H spectra of compound 2 dissolved in PBS, pH 7.0, 25°C, recorded at a concentration of 0.2 mM (**A**) and 2 mM (**B**).

Nevertheless, we succeeded in acquiring STD spectra of good quality that confirmed compound **2** ability to interact with KdsD (Figure 5).

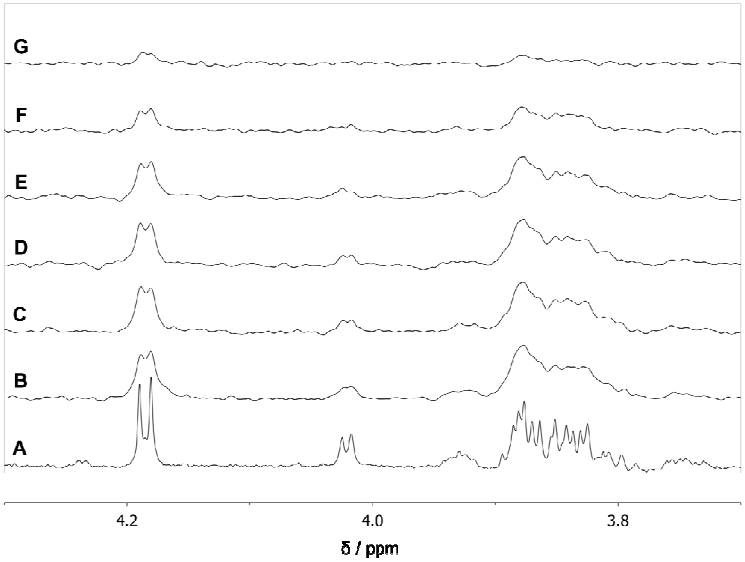


Figure 5. **A)** ^{1}H NMR spectrum of a solution of compound **2** (0.2 mM); **B-G)** ^{1}H NMR spectrum; **D-H)** STD-NMR spectra of the of the mixture containing compound **2** (0.2 mM) and KdsD (30 uM) acquired with different saturation times (**B**, 3.0 s; **C**, 2.0 s; **D**, 1.5 s; **E**, 1.0 s; **F**, 0.6 s; **G**, 0.3 s) and a saturation frequency of 7.9 ppm. All samples were dissolved in PBS, pH 7.0, 25°C.

The binding of compounds **8** was tested by STD-NMR, working on a mixture containing both the anomers at a 1:1 molar ratio; in this way, we were able to compare directly their affinity for KdsD, on the basis of the relative STD intensities. The ^{1}H NMR spectrum of a mixture containing the α and β anomers and the enzyme dissolved in PBS, pH 7.0, 37°C, is depicted in Figure 6C. This spectrum was acquired about 3 h after incubation of **8a** and **8b** with the protein and, as expected as a consequence of position 1 glycosilation, it clearly shows that they are not KdsD substrates, as their isomerization to the corresponding ketoses did not occur. Figure 7 reports the STD spectra recorded on the same mixture and indicates that both the glycosides are KdsD ligands. In addition, it indicates that the β anomer has

lower affinity for the protein, as its STD signals are less intense than those of α anomer.

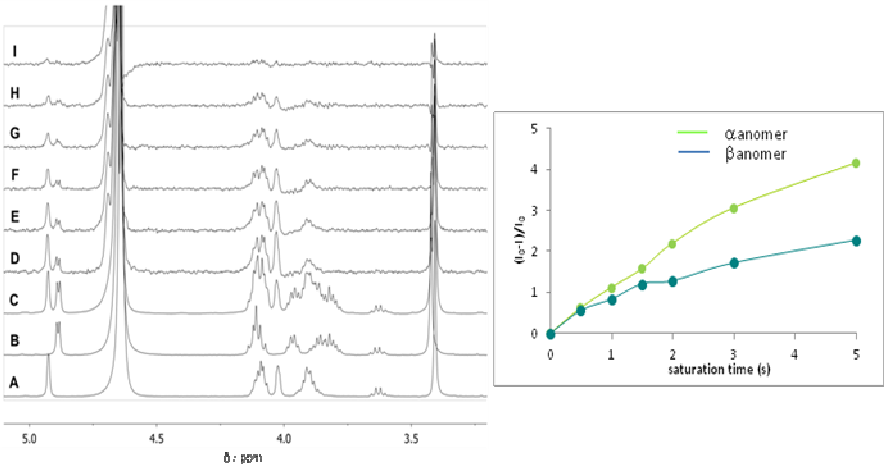


Figure 6. Left panel: **A)** ^{1}H NMR spectrum of a solution of compound 8a (3 mM); **B)** ^{1}H NMR spectrum of a solution of compound 8b (3 mM); **C)** ^{1}H NMR spectrum of a mixture containing of compound 9 (3 mM), compound 10 (3 mM) and KdsD (30 uM); **D-I)** STD-NMR spectra of the a mixture containing of compound 9 (3 mM), compound 10 (3 mM) and KdsD (30 uM) acquired with different saturation times (**D**, 3.0 s; **E**, 2.0 s; **F**, 1.5 s; **G**, 1.0 s; **H**, 0.6 s; **I**, 0.3 s) and a saturation frequency of 7.9 ppm. All samples were dissolved in PBS, pH 7.0, 37°C.

To verify if the recognition of compounds **8a** and **8b** occurred at KdsD binding site, a competitive STD with natural substrates A5P and Ru5P was performed. In particular, the competitive STD experiment was performed on a mixture containing KdsD (30 uM), the natural substrates A5P and Ru5P (total concentration 5 mM) and compounds **8a** and **8b** (3 mM each). Signals of **8a** and **8b** are almost absent in the competitive STD spectrum (Figure 7F), while natural substrates resonances appear clearly, indicating the glycosides are displace by A5P and Ru5P.This evidence demonstrated that molecules **8a** and **8b** bind the enzyme in the active site, also if with a lower affinity if compared with natural substrates.

These data suggest that probably KdsD is able to recognize the natural substrates also in hemiacetalic form; nevertheless, the cyclic form of the

ligand and/or the introduction of a substituent in anomeric position reduce dramatically the affinity.

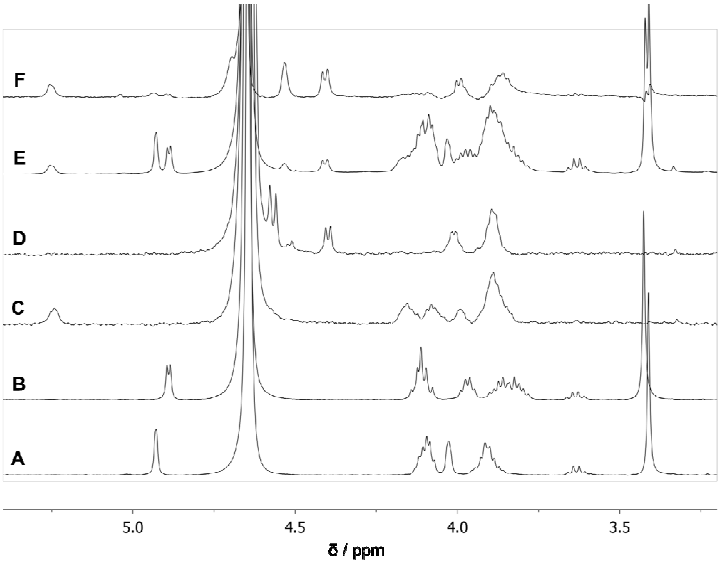


Figure 7. **A)** ^{1}H NMR spectrum of a solution of compound **8a** (3 mM); **B)** ^{1}H NMR spectrum of a solution of compound **8b** (3 mM); **C)** ^{1}H NMR spectrum of a solution of A5P (3 mM); **D)** ^{1}H NMR spectrum of a solution of Ru5P (3 mM); **E)** ^{1}H NMR spectrum of a mixture containing of compound 8a (3 mM), compound 8b (3 mM), A5P and Ru5P at equilibrium (total concentration 5 mM) and KdsD (30 uM); **F)** STD-NMR spectra acquired on the same mixture of spectrum E with a saturation time of 2.0 s and a saturation frequency of 7.9 ppm. All samples were dissolved in PBS, pH 7.0, 37°C.

Compounds **3**, **5**, **6**, **7**, **9a** and **9b** were tested as KdsD ligands according to the same experimental approach. STD data indicated that these compounds are not recognized by the enzyme, as none of the compounds resonances appear in the STD spectra acquired in the presence of KdsD (Figure 8). Thus, the proposed acidic substituents do not mimic properly the phosphate group present in the natural substrates.

Interestingly compound **4** resulted to be able to be bound and processed by the enzyme generating the ketose isomer (Figure 9 A-B). When tested in the presence of the natural substrates, it was observed that compound 4 resonances are almost absent in the spectrum, suggesting that also in this case the analog has a lower affinity to KdsD, compared to natural substrates (Figure 9 E).

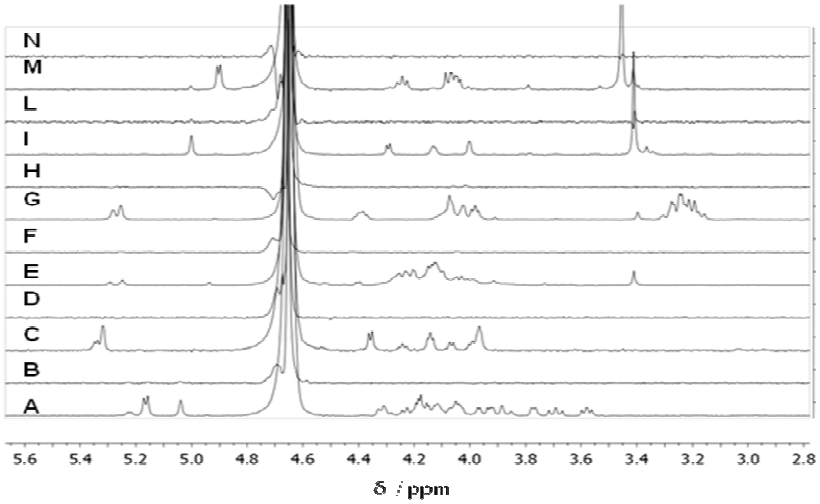


Figure 8. ^{1}H (**A**) and STD-NMR (**B**) spectra of a solution of compound **3** in the presence of KdsD; ^{1}H (**C**) and STD-NMR (**D**) spectra of a solution of compound **5** in the presence of KdsD; ^{1}H (**E**) and STD-NMR (**F**) spectra of a solution of compound 6 in the presence of KdsD; ^{1}H (**G**) and STD-NMR (**H**) spectra of a solution of compound **7** in the presence of KdsD; ^{1}H (**I**) and STD-NMR (**L**) spectra of a solution of compound **9a** in the presence of KdsD; ^{1}H (**M**) and STD-NMR (**N**) spectra of a solution of compound **9b** in the presence of KdsD. All the STD-NMR spectra were acquired with a saturation time of 3.0 s and a saturation frequency of 7.9 ppm. All samples were dissolved in PBS, pH 7.0, 37°C; compounds are 3mM, whereas KdsD is 30uM in each sample.

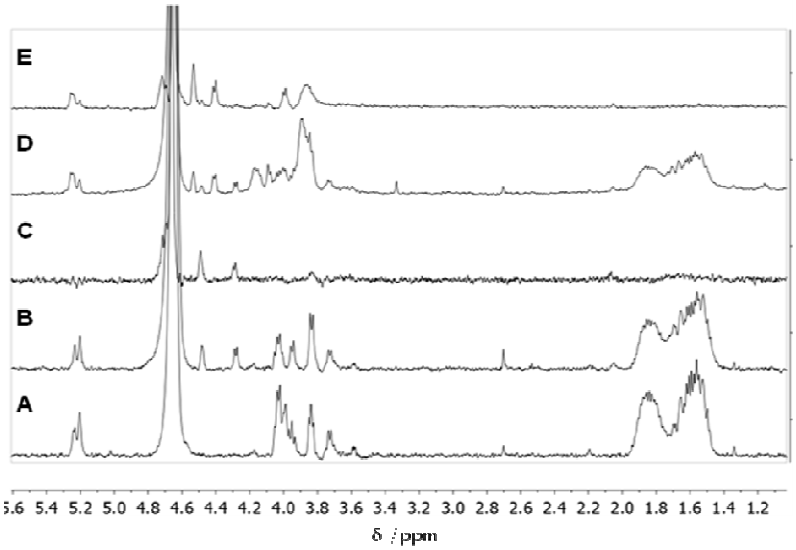


Figure 9. **A)** ^{1}H NMR spectrum of a solution of compound 4 (3 mM); ^{1}H (**B**) and STD-NMR (**C**) spectra of a solution of compound **4** in the presence of KdsD 30uM; ^{1}H (**D**) and STD-NMR (**E**) spectrum of a solution of **4** (3 mM) with natural substrates (total concentration 3 mM); STD-NMR spectra were acquired with a saturation time of 2.0 s and a saturation frequency of 7.9 ppm. All samples were dissolved in PBS, pH 7.0, 37°C.

To better investigate the properties of compounds **8a**, **8b**, **9a**, **9b** we tested their ability to inhibit the growth of *P. aeruginosa* in liquid culture. As Pa-KdsD shares high level of sequence similarity with its *E. coli* counterpart, we tested an *E. coli* mutant strain (AS19) that is more permeable to many antibiotics [4] in order to address the extent to which the Gram-negative bacteria cell barrier limits the action of the compounds. In addition, we also assayed the effect of the compounds on growth of wild type strain of *E. coli* (BW25113) in the presence or absence of glucose-6-phosphate (G6P), as it is known that G6P induces the hexose phosphate transport system (*uhp*) in *E. coli* and A5P is a high-affinity, though non-inducing, substrate of this system.[5]

As shown in Table 1, none of the compounds appeared to exert inhibitory growth effect neither on *P. aeruginosa* nor on wild type *E. coli* strain at the concentrations tested. Epitope mapping studies revealed that compounds **8a** and **8b** interact with Pa-KdsD and that this interaction occurs at the active site of the enzyme, however their binding is weaker than the natural substrates. These results are in line with the slight inhibitory activity observed *in vivo* on bacterial cultures of the permeable mutant AS19, where the minimal inhibitory concentration (MIC) of each anomers was 250 μM. Surprisingly, also compounds **9a** and **9b,** that were not recognized by the enzyme in the STD-NMR experiments, led to growth arrest of AS19 cultures with a minimal inhibitory concentration (MIC) of 250 μM. This suggests that compounds **9a** and **9b** act as weak inhibitors, possibly by interacting with the Kdo 8P synthase, the enzyme acting downstream KdsD in the Kdo biosynthetic pathway. However, an intact outer membrane normally prevents the entry of these compounds into the cell as no inhibitory activity is observed against wild type *E. coli* and *P. aeruginosa* strains. This suggests that rounds of optimization are still required to improve cellular uptake of these compounds.

On the basis of the information obtained, other derivatives are being prepared in our laboratory.

Table 1. MIC determination of compounds 8-9.

	MIC, mM			
	Compound:			
Strain	**8a**	**8b**	**9a**	**9b**
E. coli BW25113	>2	>2	>2	>2
E. coli BW25113 +G6P	>2	>2	>2	>2
E. coli AS19	0.25	0.25	0.25	0.25
P. aeruginosa PA01	>2	>2	>2	>2

Experimental Section

Compounds 1 and 2 were kindly provided by the group of R.W. Woodard, whereas the compounds 3-9 here described were synthesized in our laboratory and purified by standard flash chromatography techniques. The ligand resonances were assigned by ^{1}H, ^{1}H,^{1}H-COSY and ^{1}H,^{13}C-HSQC NMR spectroscopy.[102]

NMR spectroscopy experiments

All of the experiments were recorded with a Varian Mercury (400 MHz) or a Bruker Advance III (600 MHz) instrument with a cryoprobe. The ligand resonances were assigned by ^{1}H,^{1}H COSY and ^{1}H,^{13}C-HSQC NMR spectroscopy. A basic 1D-STD sequence was used with the on-resonance frequency of 7.9 ppm and the off-resonance frequency of 40 ppm. A train of Gaussian-shaped pulses of 50 ms each was employed, with total saturation times of the protein envelope ranging from 5.0 to 0.3 s. The total saturation time was adjusted by the number of shaped pulses. A $T_{1\rho}$ filter of 2 ms was employed to eliminate the background signals from the protein. All of the samples were dissolved in deuterated PBS (pH 7.0) at 25 or 37° C. Total

[102] Manuscript in preparation

sample volumes were 0.55 mL. The on- and off-resonance spectra were acquired simultaneously with the same number of scans. The STD spectrum was obtained by subtraction of the on-resonance spectrum from the off-resonance spectrum. Subtraction was performed by phase cycling to minimise artefacts arising from magnet and temperature instabilities. Reference experiments with samples containing only the free compounds tested were performed under the same experimental conditions to verify true ligand binding. The effects observed in the presence of the protein were due to true saturation transfer, because no signals were present in the STD spectra obtained in the reference experiments, except for residues from HDO, indicating that artefacts from the subtraction of compound signals were negligible.

Expression and purification of Pa-KdsD:

The *P. aeruginosa* KdsD (Pa-KdsD, PA4457) was expressed by pQE31/Pa-KdsD plasmid that encodes Pa-KdsD with an N-terminal hexahistidine tag[2]. Cultures of *E. coli* M15/pREP4 harboring pQE31/Pa-KdsD plasmid were grown overnight in LD broth supplemented with ampicillin (100 µg mL^{-1}) and kanamycin (25 µg mL^{-1}), diluted 1:100 in fresh medium (1 L) and grown up to mid-logarithmic phase (OD_{600}, 0.6) at 37 °C with shaking. Induction of Pa-KdsDwas then carried out for 4 h with IPTG (0.5 mM). Cells were harvested, washed with sodium phosphate (20 mM, pH 7.0) and re-suspended in sodium phosphate (50 mM, pH 8.0), NaCl (300 mM), imidazole (10 mM) supplemented with DNase (0, 1 mg mL^{-1}) and phenylmethylsulfonyl fluoride (1 mM). After treatment with lysozyme (1 mg mL^{-1}) for 30 min, cells were disrupted by a single cycle through a Cell Disruptor (One Shot Model by Constant Systems LTD) at 25,000 psi and unbroken cells were removed by centrifugation at 39,000g for 30 min. The

supernatant was loaded at a flow rate of 0.5 mL min^{-1} onto a Ni-NTA Agarose (Qiagen) column (bed volume 5 mL) pre-equilibrated with sodium phosphate (50 mM, pH 8.0) and NaCl (300 mM). The column was washed with 10 volumes of sodium phosphate (50 mM, pH 8.0), NaCl (300 mM) and imidazole (20mM) and the enzyme was eluted with a 5 mL step-wise imidazole gradient (100 to 300 mM) in the same buffer. Fractions (5 mL) were collected and those displaying the highest purity (>95%, as assessed by SDS-PAGE analysis) were dialyzed in sodium phosphate (50 mM, pH 8.0).The isomerase activity of the protein was assessed as described previously[2]. Protein aliquots (600 µg each) were lyophilized and stored at - 80 °C.

MIC determination

Minimal inhibitory concentrations (MIC) were determined following standard protocols for testing susceptibility to antibiotic agents [6]. The bacterial strains used were *Pseudomonas aeruginosa* PAO1 (ATCC 15692), *Escherichia coli* K12 wild type strain BW25113 [7], and the permeable *E. coli* mutant strain AS19 [4]. The procedure for a broth microdilution test was followed. Briefly, 96-well microtiter plates were prepared with compound concentrations ranging from 2mM to 2µM, in triplicates. For *E. coli* BW25113 the procedure was performed in presence or absence of 10 µM glucose 6-phosphate. The assayed strains were grown on Mueller-Hinton broth (MHB) and inoculums of approximately 5×10^5 cfu mL^{-1} were added to each well. The microtiter plates were incubated for 18-20 h at 37°C and the MIC of each compound determined. MIC was defined as the lowest concentration inhibiting bacterial growth.

References

[1] Airoldi C., Sommaruga S., Merlo S., Sperandeo P., Cipolla L., Polissi A., Nicotra F. Targeting Bacterial Membranes: NMR characterization of substrate recognition and

binding requirements of D-arabinose 5-phosphate isomerase, *Chem. Eur. J.*, 2010, 16, 1897 – 1902
[2] Airoldi C., Sommaruga S., Merlo S., Sperandeo P., Cipolla L., Polissi A., Nicotra F. Targeting Bacterial Membranes: identification and NMR Characterization of Substrate Recognition and Binding of *P. aeruginosa* D-Arabinose-5P Isomerase, *ChemBioChem*, 2011, 12, 719-727.
[3] Yep A., Sorenson R.J., Wilson M.R., Showalter H.D.H, Larsen S.D., Keller P.K., Woodard R.W., Enediol mimics as inhibitors of the D-arabinose 5-phosphate isomerase (KdsD) from *Francisella tularensis*, Bioorganic and Medicinal chemistry Letters, 2011, 21, 2679-2682.
[4] Sekiguchi, M. & Iida, S. Mutants of *Escherichia coli* permeable to actinomycin. Proc. Natl. Acad. Sci. USA, 1967, 58, 2315–2320.
[5] Meredith TC, Woodard RW. Identification of GutQ from Escherichia coli as a D-arabinose 5-phosphate isomerase. J Bacteriol. 2005, 187:6936-42.
[6] Wiegand I., Hilpert K., Hancock R. E. Agar and broth dilution methods to determine the minimal inhibitory concentration (MIC) of antimicrobial substances. Na.t Protoc. 2008, 3, 163.
[7] Datsenko K.A., Warren B.L. One-step inactivation of chromosomal genes in E.coli K-12 using PCR products. Proc Natl Acad Sci. 2000, 97, 6640-6645

NMR characterization of glycofused tricycles as ligands for Prion Protein 106-126: trying to develop new tools for the diagnosis and therapy of neurodegenerative disorders.

Introduction

Despite the efforts in studying PrP aggregation and toxicity, prion diseases are invariably fatal, and no viable treatments for these devastating disorders are currently available.[1] Many small molecules that are able to bind these peptides and inhibit their aggregation are already known; most of them are natural compounds bearing aromatic moieties. Among them tetracyclines have been shown to bind to PrP 106-126, to inhibit PrP aggregation into amyloid fibrils, and to prevent neuronal death. However, they are characterized by a low solubility and low chemical stability in physiological conditions and show other pharmacological activities not directly related to amyloidosis.

Trying to overcome these limitations, in our laboratory a small panel of stable and solubile glycofused aromatic tricyclic compounds have been synthesized[2] (Figure 1) and tested by NMR for their ability to interact with PrP oligomers.[103]

These molecules were originally designed as Aβ-peptide ligands[2], but, as it was reported that often molecules able to inhibit aggregation in one amyloidogenic process are also active against other similar pathologies, we decided to study their behavior against PrP106-126.

[103] Manuscript in preparation

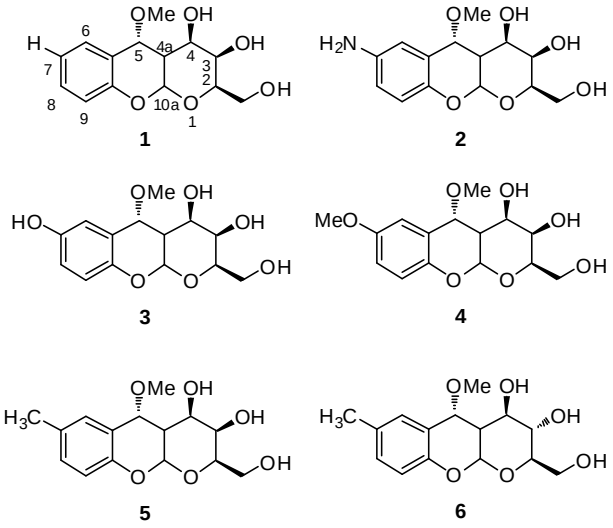


Figure 1. Structures of the tricyclic compounds tested for their interaction with PrP106-126.

Results and discussion

These tricyclines are characterized by a glyco-fused benzopyran structure, that maintains an aromatic moiety. By performing NMR interaction studies, we demonstrated that these molecules are good ligands of PrP106-126 aggregates (Figure 2).

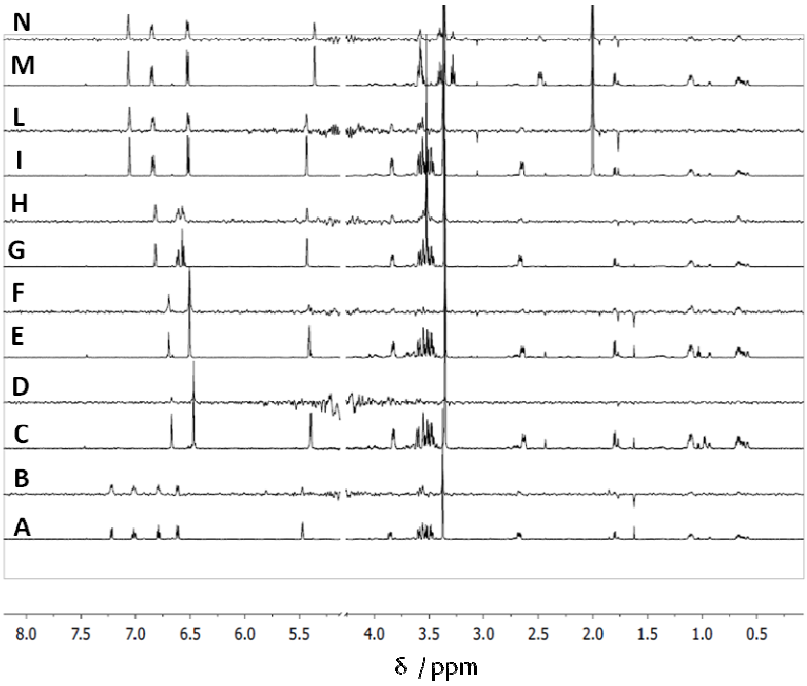


Figure 2. ^{1}H spectra of compounds **1**, **2**, **3**, **4**, **5**, **6** dissolved in deuterated PBS at 25°C (**A**, compound **1**; **C**, compound **2**; **E**, compound **3**; **G**, compound **4**; **I** compound **5, M** compound **6**) and 1D-STD spectra of the same mixtures dissolved in deuterated PBS at 25°C, containing PrP106-126 (95 uM) and one of the tested compounds (1 mM) (**B**,compound **1**; **D**, compound **2**; **F**, compound **3**; **H**, compound **4, L**, compound **5, N**, compound **6**).

In particular, STD-NMR spectra indicated that the sugar moiety is less involved in the binding process, as its signals are less intense in comparison with the aromatic protons (Figure 3), and therefore it can be further functionalized, in order to improve the binding properties of the molecules.

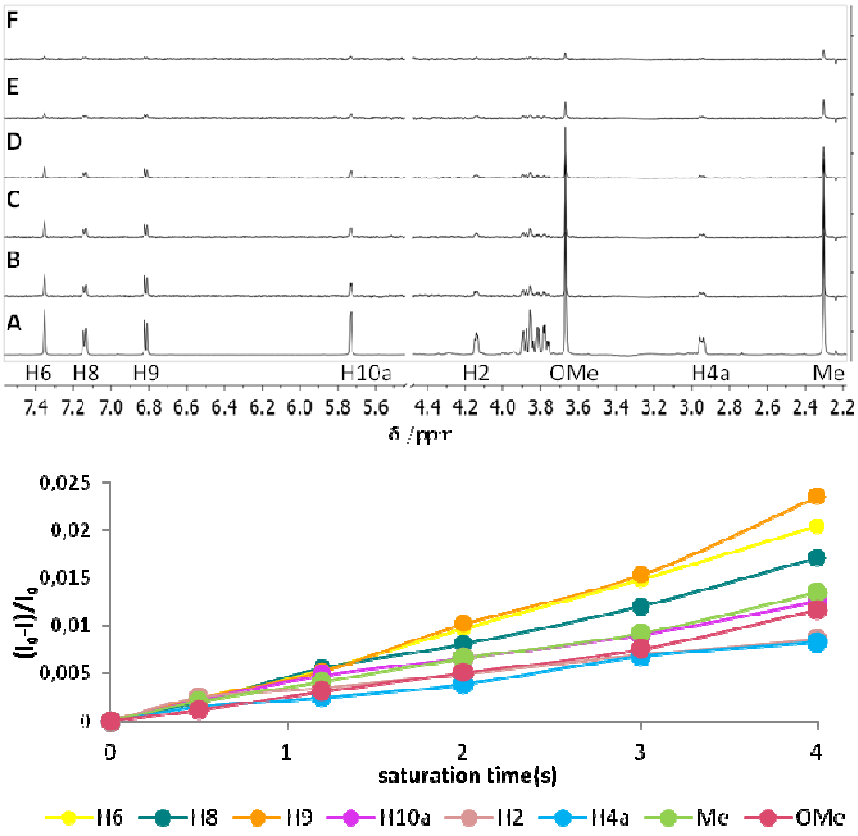


Figure 3. Top panel: A) 1H-NMR spectrum of compound 5 (1mM) in the presence of PrP106-126 (95uM); B-F) STD-NMR spectra of the mixture recorded with different saturation time (4s (B), 3s (C), 2s (D), 1.2s (E), 0.5s (F)); Bottom panel: Fractional STD effect for each non-overlapping signal of compound 5, calculated by $(I_0\text{-}I)/I_0$, in which $(I_0\text{-}I)$ is the peak intensity in the STD spectrum and I_0 is the peak intensity in the off-resonance spectrum.

By performing a competitive STD, in which all the tricyclines were mixed together with the PrP, we succeeded in ranking the molecules on the basis of their affinity to PrP (Figure 4). It was found that lower the polarity of the substituents on the aromatic ring, higher the affinity to PrP, confirming the importance of their hydrophobic character to allow for a strong interaction with PrP. In fact, compounds whose aromatic rings are substituted with a

methyl group (**5**, **6**) are the ligands with the highest affinities for PrP aggregates, followed by those presenting an *O*-methyl group as a substituent (**4**), and by **1**, with no aromatic substituent.

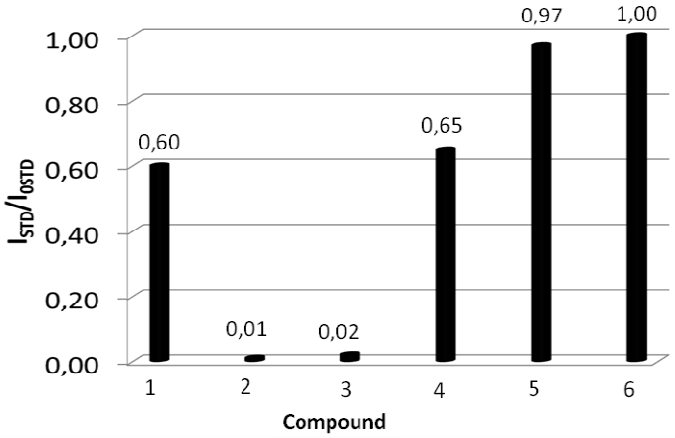


Figure 4. Fractional STD effects calculated for compounds **1**, **2**, **3**, **4**, **5**, **6**. These values are proportional to compound affinity for PrP106-126 aggregates.

It was already reported that these molecules have the same conformation and therefore the differences in affinity for PrP106-126 peptide are not consequence of conformational differences, but they are just determined by the difference in the polarity.[2]

Then binding of the best ligands (compounds **5**, **6**) was confirmed by the acquisition of a transfer-NOESY. The change in the sign of the cross-peaks of the test molecule, from positive, in the absence of PrP106-126, to negative, in the presence of PrP106-126, reflects an increase of its effective rotational motion correlation time, and supports its binding to a large molecular entity, here represented by the PrP106-126 aggregates (Figure 5).

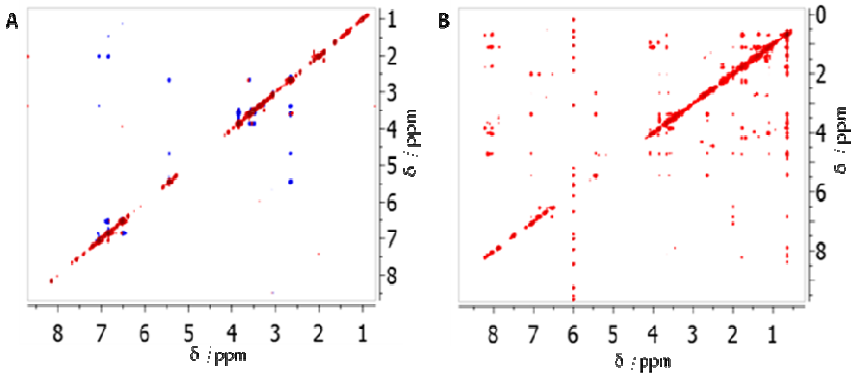


Figure 5. A) NOESY spectrum of compound 5 alone (1mM); B) tr-NOESY of compound 5 (1mM) in the presence of PrP (0.5mM).

It was also demonstrated that these molecules do not present the undesired antibacterial activity of tetracycline. MIC results for the reference tetracycline and compounds **1**-**6** on different bacterial strains are reported in the following table.

Table 1. MIC results on tetracycline and compounds 1-6.

	MIC, μM						
Strain	Tetracycline	compound 1	compound 2	compound 3	compound 4	compound 5	compound 6
E. coli MG1655	31.3	>500	>500	>500	>500	>500	>500
E. coli AS19	1	>500	>500	>500	>500	>500	>500
P. aeruginosa PA01	250	>500	>500	>500	>500	>500	>500
S. aureus	7.8	>500	>500	>500	>500	>500	>500
S. epidermidis	7.8	>500	>500	>500	>500	>500	>500

Trying to develop a diagnostic tool, another tricyclic molecule was generated (Figure 6): it is characterized by the functionalization of the sugar moiety with a short chain of polyethylene glycol (PEG) to which coumarin has been linked.

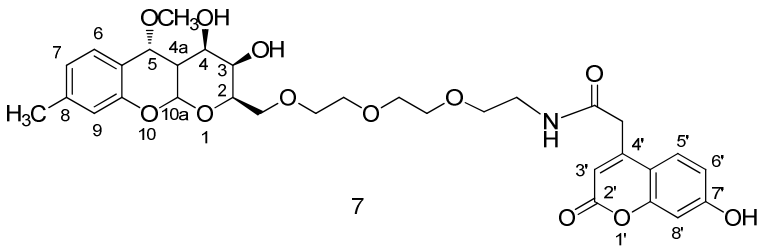


Figure 6. Tricyclic fluorescent-derivative.

Coumarin is a fluorescent probe that can be used in order to stain PrP fibrils, making possible to visualize the plaques. Moreover coumarin was selected as fluorophore for tricycline derivatization in order to modulate the hydrophilic properties of the molecule, maintaining the molecular weight low enough for a diffusion mechanism. Indeed, in order to be useful for diagnostic applications, this compound should perform its action within the brain, therefore it has to be able to cross the blood brain barrier (BBB). The BBB is formed by the complex tight junctions between the endothelial cells of the brain capillaries and their low endocytic activity. This results in the capillary wall that behaves as a continuous lipid bilayer and prevents the passage of polar and lipid insoluble substances. Unlike the most of the tissue, the BBB prevents the paracellular passage, thus the small molecule should pass exploiting a transcellular mechanism. It is, therefore, the major obstacle to drugs that may combat diseases affecting the central nervous system.[5] Only few compounds with the correct hydrophilic/lypophilic balance have a greater chance to overcome the BBB through a diffusion mechanism. The passive diffusion of a drug depends on its blood/brain concentration gradient and its lipid solubility, but it is inversely related to its degree of ionization and its molecular weight. Factors other than lipophilicity and molecular weight also modulate the transport of a drug across the BBB. Reducing the relative number of polar groups increases the

transfer of a drug across the BBB. Evaluation of drug transport to the brain *in vitro* has usually been carried out by studying the transport of individual molecules across endothelial cell monolayers. Currently available *in vitro* models for BBB allow to evaluate quickly and in a reproducible way the predictive *in vivo* permeability of compounds and drugs under development.[4,5] Fluorescence mesurements carried out with the fluorescent tricycline derivative indicate that its favorable physical–chemical properties, in terms of balanced hydrophylicity/lipophylicity, allow this compound to permeate through the BBB, most probably exploiting a diffusion mechanism. Moreover we verified, through STD-NMR, that this compound effectively retains the binding properties of the tricyclines (Figure 7); all the other biological tests are in progress.

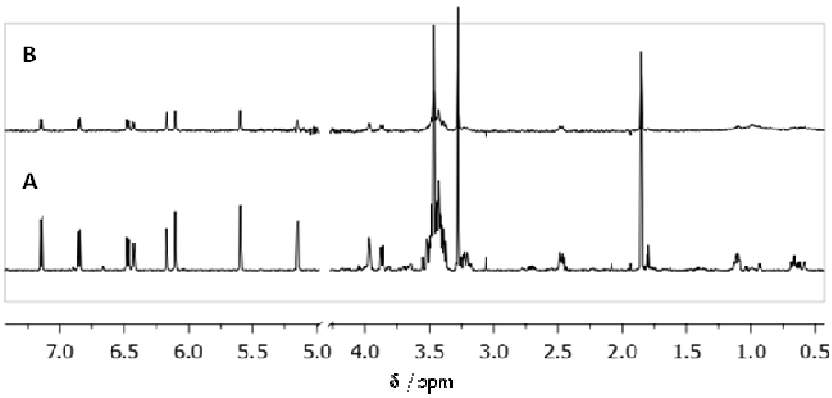


Figure 7. A) ^{1}H-NMR spectrum of the compound 7 (1mM) in the presence of PrP106-126 (95uM); B) STD-NMR spectrum of the mixture.

Experimental Section

All the compounds here described were synthesized in our laboratory and purified by standard flash chromatography techniques. The ligand resonances were assigned by ^{1}H, ^{1}H,^{1}H-COSY and ^{1}H,^{13}C-HSQC NMR spectroscopy.[2]

PrP(106-126) was synthesized and provided by the research group of Mario Salmona, at the Mario Negri Institute for Pharmacological Research (Milan).

NMR spectroscopy experiments

All the experiments were recorded with a Bruker Advance III (600 MHz) instrument equipped with a cryoprobe. A basic 1D-STD sequence was used with the on-resonance frequency of -0.5 ppm and the off-resonance frequency of 40 ppm. A train of Gaussian-shaped pulses of 50 ms each was employed, with total saturation times of the protein envelope ranging from 4.0 to 0.5 s. The total saturation time was adjusted by the number of shaped pulses. All of the samples were dissolved in deuterated PBS (pH 7.4) at 5°C. Total sample volumes were 600 uL. The on- and off-resonance spectra were acquired simultaneously with the same number of scans. The STD spectrum was obtained by subtraction of the on-resonance spectrum from the off-resonance spectrum. Reference experiments with samples containing only the free compounds tested were performed under the same experimental conditions to verify true ligand binding. The effects observed in the presence of the protein were due to true saturation transfer, because no signals were present in the STD spectra obtained in the reference experiments, except for residues from HDO, indicating that artefacts from the subtraction of compound signals were negligible.

MIC determination

Minimal inhibitory concentrations (MIC) were determined following standard protocols for testing susceptibility to antibiotic agents.[6]

The bacterial strains used were: *Pseudomonas aeruginosa* PAO1 (Gram negative), *Escherichia coli* K12 wild type strain MG1655 (Gram negative), the permeable *E. coli* mutant strain AS19[7] (Gram negative), *Staphylococcus aureus* (Gram positive) and *Staphylococcus epidermidis*

(Gram positive). The procedure for a broth microdilution test was followed. Briefly, 96-well microtiter plates were prepared with compound concentrations ranging from 500 to 0.5 µM, in triplicates. As a positive control tetracycline hydrochloride (SIGMA) was used at the same concentrations. The assayed strains were grown in Mueller-Hinton broth (MHB) and inoculums corresponding to OD_{600} of 0.05 were added to each well. The microtiter plates were incubated for 18-20 h at 37°C and the MIC of each compound determined. MIC was defined as the lowest concentration inhibiting bacterial growth.

References

[1] J. Med. Chem. 2011, 54, 1010–1021
[2] *Chem.Commun.*, 2011, 47, 10266-10268.
[3] *Vascul Pharmacol*., 38, 6, 2002, 349-54.
[4] Nanomedicine, 2011, 7(5), 551-559
[5] Adv. Drug Deliver Rev., 1999, 36, 165–178.
[6] Nat Protoc.,2008, 3, 163.
[7] Proc. Natl. Acad. Sci. USA, 1967, 58, 2315–2320.

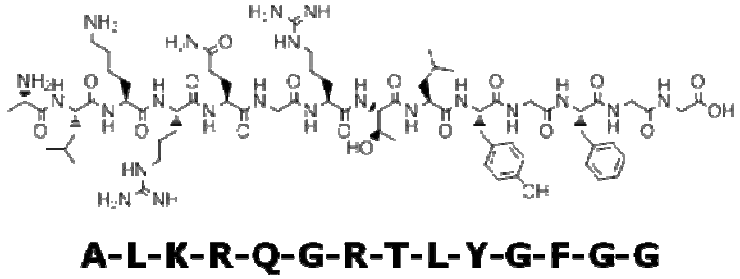
Osteogenic growth peptide interaction with α-2-Macroglobulin: preliminary NMR studies functional to biomaterial generation

Introduction

Biomaterials were traditionally defined as materials used in medical devices and have been used since antiquity, but nowadays their degree of sophistication has increased significantly, as advances in molecular biology, genomics and proteomics affected the way in which biomaterials are designed and used. In particular, specific molecules recognized for their importance in clinically relevant processes have been incorporated into the materials as bioactive component, mimicking the natural cues responsible for cell adhesion, tissue regeneration, and therefore creating a smart material that can interact with the biology of the host.[1]

For this purpose it is necessary to choose natural ligands suitable for the material decoration and to verify that, after immobilisation on the surface, ligands retain the proper conformation and therefore the ability to interact with the specific binding partners.

Among the different cues, involved in cell growth and diffrentiation, we identified the Osteogenic Growth Peptide (OGP) as an interesting molecule to be studied. It is a highly conserved, naturally occurring tetradecapeptide (Figure 1) discovered in the early 1990s that has attracted considerable interest as a bone anabolic agent and a hematopoietic stimulator, as it exerts regulatory effects on bone and bone marrow and shows mitogenic effects on fibroblasts.



A-L-K-R-Q-G-R-T-L-Y-G-F-G-G

Figure 1. OGP sequence.

It is physiologically present in the blood circulation, mainly in the form of an OGP-OGP binding protein (OGPBP) complex.[2,3,4] α-2-Macroglobulin (α2M) is a plasma OGPBP of 720 kDa and isan important modulator of the OGP action: it maintains large reservoirs of inactive OGP protected from proteolytic degradation and provides a mechanism for the control of peptide availability to its target cells.[5] α2M exists in two states: the native form enhances the immediate availability of OGP to its target cells, whereas activated α2M may participate in the removal of OGP from the system.[6] Following its dissociation from the complex with α2M, OGP seems to be proteolytically cleaved, generating the C-terminal pentapeptide that is the minimal OGP-derived sequence that retains the full OGP-like biological potency, and appears to lack OGPBP binding properties. [5,7] The putative receptor(s) through which OGP exerts its activity is(are) still unknown[5] and there are no structural data on the peptide conformation in physiological conditions and in the presence of α2M. Considering the key role of α2M in regulating OGP activity, we are currently exploiting high resolution NMR techniques in order to investigate the formation of OGP – α2M complex at molecular level. Moreover, due to peptide osteogenic properties, biomaterials for tissue engineering will be decorated with OGP/OGP mimetics and, as soon as the peptide will be linked to a suitable

scaffold, conformational and binding studies on the generated material will be performed.

Results and discussion

In physiological conditions (PBS, pH 7.4) the OGP-amide NMR signals are very broad, suggesting that the peptide does not possess a rigid conformation. In the same conditions but in the presence of α2M some amide signals are shifted and become sharper; moreover, there are some other shifts in the peptide resonances. Chemical shift variations mainly regard the peptide backbone, suggesting that in the presence of α2M, OGP adopts a more defined conformation. (Figure 2-3)

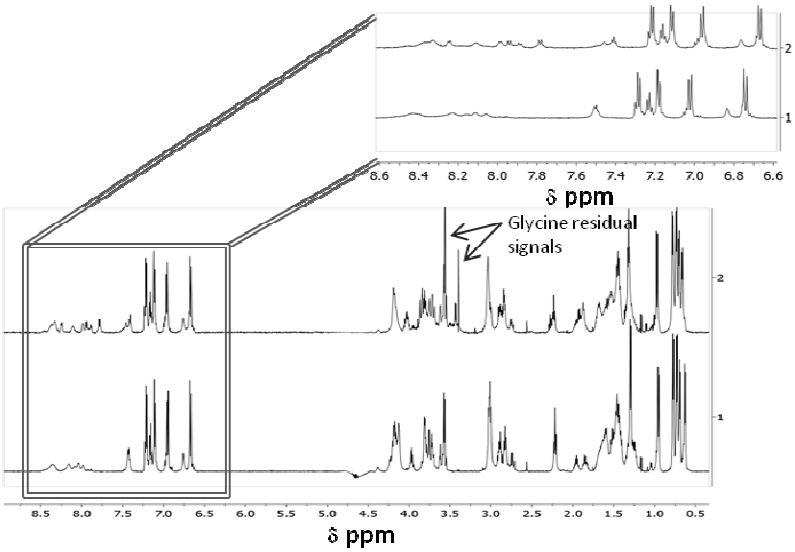


Figure 2. 1) ^{1}H NMR spectrum of OGP 1mM, PBS, pH 7.2, 25°C; 2) ^{1}H NMR spectrum of OGP 1mM in the presence of 5uM a2M, PBS, pH 7.2, 25°C. (*The protein was purchased lyophilized in a buffer containing 0.1 M glycine. It was dialyzed o.n.)

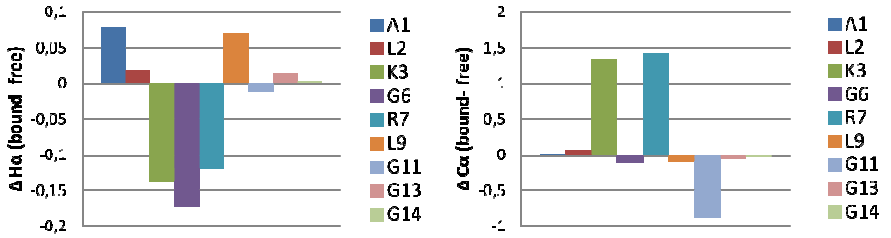


Figure 3. Peptide shifts observed in the ^{1}H-^{13}CHSQC-NMR spectrum in the presence of α2M. Left panel: Hα shifts calculated as difference between bound and free OGP; right panel: Cα shifts calculated as difference between bound and free OGP.

After having assigned all the peptide resonances, we performed interaction studies in the presence of α2M. In particular, we acquired STD-NMR[8] and Water-Ligand Observed via Gradient Spectroscopy (WaterLOGSY)[9] experiments.

STD spectra recorded on the ligand : protein mixture show that peptide regions involved in the binding are mainly the methyl groups of Leu, Thr and Ala and the aromatic residues Phe and Tyr. (Figure 4)

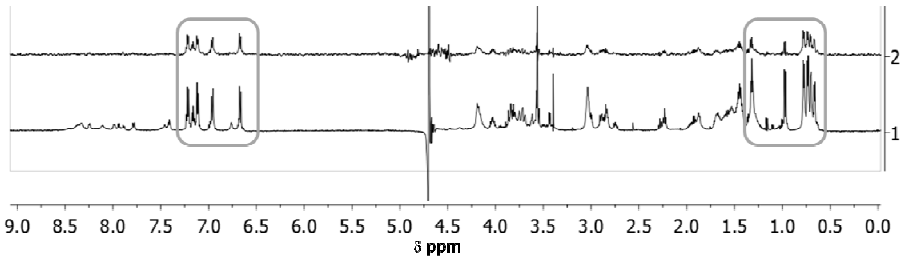


Figure 4. 1) ^{1}H-NMR spectrum of OGP in the presence of α2M; 2) STD-NMR spectrum recorded with a saturation time of 2s, and on-resonance frequency of -1ppm. The spectra were recorded on the same sample, containing 1mM OGP and 5uM α2M, PBS, pH 7.2, 25°C.

WaterLOGSY experiments confirmed these results: in fact the intensities of the same signals change from negative to positive in the presence of α2M, suggesting that the corresponding residues are directly involved in the

interaction with α2M and that interactions with bulk water are precluded. (Figure 5).

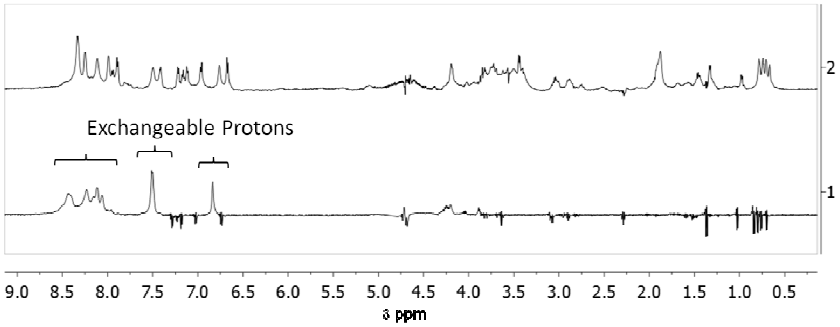


Figure 5. 1) WaterLOGSY spectrum of free OGP; 2) WaterLOGSY spectrum of OGP in the presence of α2M. Spectra were recorded with 1s of sat. time; samples were dissolved in PBS, pH 7.2, 25°C.

Due to the importance of the C-terminal pentapeptide for the biological activity, we decided to study the two OGP fragments one apart from the other. Both the OGP fragments do not seem to have a defined structure in solution.

As expected, α2M binds OGP(1-9), confirming the importance of the N-terminal portion of OGP in OGPBP binding. Surprisingly also OGP(10-14) shows STD signals when it is in the presence of α2M; its binding to the OGPBP was confirmed also by WaterLOGSY experiments (Figure 6).

In order to verify the specificity of this binding event, we performed competitive STD experiments: STD spectra were recorded on a mixture containing a OGP-fragment (OGP(1-9) or OGP(10-14)), the full-length OGP and α2M. STD intensities of OGP(1-9) resonances in the presence of the full-length OGP were lower than those obtained for the peptide alone, indicating that the two peptide compete for the same binding site and that their affinities are comparable (Figure 7). The same experimental approach was carried out on OGP(10-14); we performed a series of competitive STD

experiments by increasing OGP concentration until a 2:1 OGP:OGP(10-14) ratio was reached. No signal decrease was observed, therefore suggesting that the smaller fragment does not compete with the full-length OGP and binds to α2M in a different binding site (Figure 8).

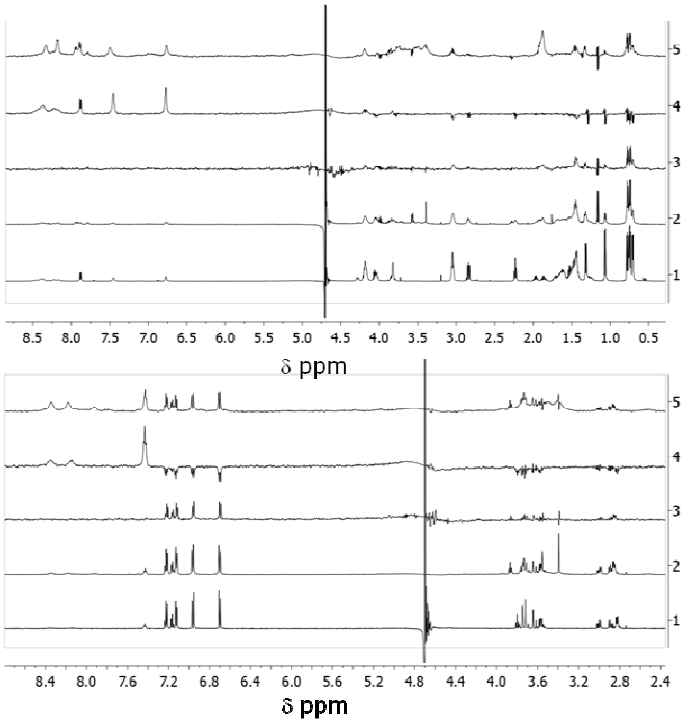


Figure 6. Top panel: 1) ^{1}H NMR spectrum of OGP(1-9) 1mM; 2) ^{1}H NMR spectrum of OGP(1-9) 1mM in the presence of 5uM a2M; 3) STD spectrum of OGP(1-9) 1mM in the presence of 5uM a2M (sat. time 2s, irr. freq. -1ppm); 4) waterLOGSY spectrum of OGP(1-9) alone (sat. time 1s); 5) waterLOGSY of OGP(1-9) 1mM in the presence of 5uM a2M (sat. time 1s); bottom panel: 1) ^{1}H NMR spectrum of OGP(10-14) 1mM; 2) ^{1}H NMR spectrum of OGP(10-14) 1mM in the presence of 5uM a2M; 3) STD spectrum of OGP(10-14) 1mM in the presence of 5uM a2M (sat. time 2s, irr. freq. -1ppm); 4) waterLOGSY spectrum of OGP(10-14) alone (sat. time 1s); 5) waterLOGSY of OGP(10-14) 1mM in the presence of 5uM a2M (sat. time 1s).
All the samples were dissolved in PBS, pH 7.2, 25°C.

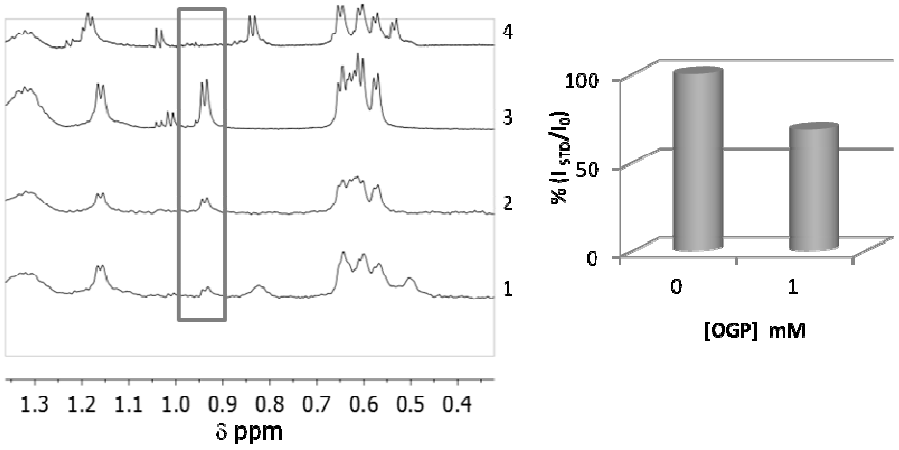


Figure 7. Left panel: zoom on the aliphatic region of the OGP ^{1}H-NMR spectrum(4), OGP(1-9) ^{1}H-NMR spectrum (3), STD spectrum of OGP(1-9) (1mM) in the presence of α2M (7uM) prior to OGP addition (2), competitive STD spectrum recorded on the same mixture after OGP addition (1mM) (1); Right panel: STD intensity decrease after the addition of OGP.

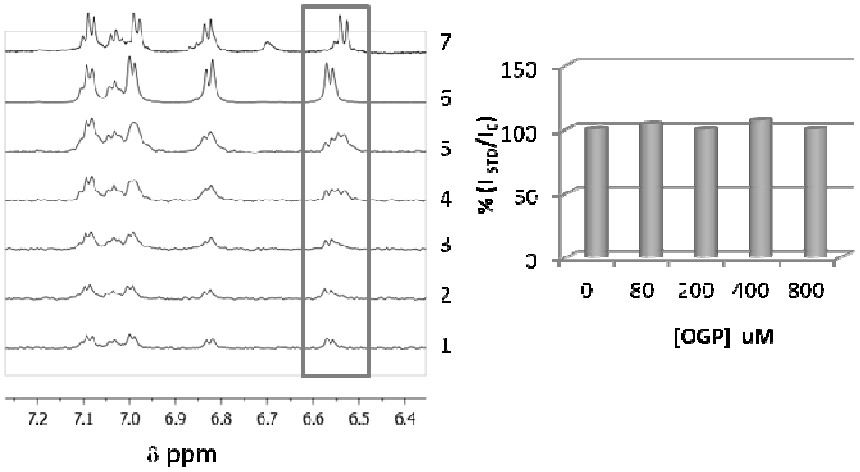


Figure 8. Left panel: zoom on the aromatic region of the OGP ^{1}H-NMR spectrum (7); OGP(10-14) ^{1}H-NMR spectrum (6); STD spectrum of OGP(10-14) (400uM) in the presence of α2M (7uM) prior to OGP addition (1); competitive STD spectra recorded on the same mixture after OGP addition (80uM, 2; 200uM, 3; 400um, 4; 800uM, 5); Right panel: STD intensity after the addition of OGP, dispaying that no significative variation of the STD intensities are evident.

We tried to investigate OGP conformation by acquiring and analyzing 2D-NOESY and tr-NOESY experiments on OGP alone and in the presence of

α2M, respectively. In compare with 2D-NOESY of the free ligand, trNOESY shows additional OGP intramolecular cross-peaks. We tried to analyze these data, using CARA[10] and CYANA[11] softwares, in order to determine OGP structure starting from NMR constraints. Unfortunately these constraints are not enough to generate a structure different from the random coil conformation observed for the free-OGP in solution. Hence, the presence of α2M somehow stabilize the peptide, but probably the intramolecular distances between protons are larger than 5Å and therefore not measurable by these experiments. This hypothesis is supported by the great differences in the ligand and the protein relative sizes : the small OGP (1523 Da) could be easily accommodated in the α2M (720 kDa) binding site without having to rearrange substantially its conformation.

As soon as OGP mimetics will be synthesized and, along with OGP, will be employed for biomaterial decoration, we will perform conformational and interaction studies at biomaterial surface.

Experimental Section

Peptides (OGP, OGP(10-14) and OGP(1-9)) were purchased from Proteogenix, whereas α2M was purchased from antibodies-online GmbH.

The lyophilized α2M was resuspended in 100 uL of milliQ-water and dialyzed over night at 4°C, using Slide-A-Lyzer®MINI Dialysis Unit (EuroClone S.p.A.), in phosphate buffer, in order to exchange the glycine buffer in which α2M was stored.

NMR spectroscopy experiments

All of the experiments were recorded with a Bruker Advance III (600 MHz) instrument equipped with a cryoprobe. The ligand resonances were assigned by $^{1}H,^{1}H$ COSY, $^{1}H,^{1}H$ TOCSY , $^{1}H,^{1}H$ NOESY or $^{1}H,^{1}H$ ROESY, $^{1}H,^{13}C$-HSQC and 2D-HSQC-TOCSY NMR experiments. WaterLOGSY experiments

were carried on with different saturation times, ranging from 1.4 to 0.7s; reference experiments with samples containing only the free peptides were performed under the same experimental conditions.

A basic 1D-STD sequence was used with the on-resonance frequency of -1.0 ppm and the off-resonance frequency of 40 ppm. A train of Gaussian-shaped pulses of 50 ms each was employed, with total saturation times of the protein envelope ranging from 3.0 to 0.35 s. The total saturation time was adjusted by the number of shaped pulses. A $T_{1\rho}$ filter of 2 ms was employed to eliminate the background signals from the protein. The on- and off-resonance spectra were acquired simultaneously with the same number of scans. The STD spectrum was obtained by subtraction of the on-resonance spectrum from the off-resonance spectrum. Reference experiments with samples containing only the free compounds tested were performed under the same experimental conditions to verify true ligand binding. The effects observed in the presence of the protein were due to true saturation transfer, because no signals were present in the STD spectra obtained in the reference experiments, except for residues from HDO, indicating that artefacts from the subtraction of compound signals were negligible.

All the samples were dissolved in PBS (pH 7.2) at 25°C or 5°C. Total sample volumes were 600 uL (for 5mm NMR-tubes) or 200uL (for 3mm NMR-tubes).

References

[1] Nature, 2009, 462, 426-432
[2] Journal of Cellular Biochemistry, 2006, 98, 1007-1020
[3] Leukemia Research, 2002, 26, 839-848
[4] The EMBO Journal, 1992, 11, 5, 1867-1873
[5] Biopolymers (Peptide Science), 2002, 66, 33-48
[6] Biochemistry, 1997, 36, 14883-14888
[7] Journal of Peptide Research, 2000, 56, 147-156

[8] Angew. Chem. Int. Ed. 1999, 38, 12, 1784-1788
[9] Journal of Biomolecular NMR, 2001, 21, 349–359
[10] http://cara.nmr-software.org/downloads/
[11] http://www.cyana.org

Journal of Materials Science and Engineering B 2 (12) (2012) 618-625

NMR Protein-Ligand Interaction Studies under Non-Homogeneous Conditions for Biomaterial Generation: A Model For Artificial Lectin-Carbohydrate Recognition

Cristina Airoldi,[1] Silvia Merlo, [1] Erika Sironi, [1] Francesco Nicotra, [1]
Jesús Jiménez-Barbero [1,2]

1. Department of Biotechnology and Biosciences, University of Milano-Bicocca, 20126, Milan, Italy
2. Centro de Investigaciones Biológicas, CSIC, 28040 Madrid, Spain

Abstract: Smart biomaterials for tissue regeneration need to incorporate molecules able to interact with specific cellular adhesion or morphogenic proteins of the extracellular matrix (ECM). NMR binding studies allow obtaining structural information essential for the comprehension of biological processes and nowadays high-resolution magic-angle-spinning (HR-MAS) NMR spectroscopy is a well-established tool for the study of heterogeneous systems. Here we present the generation of a model-system used to explore the possibility to reveal interactions between two molecular entities, one of which linked to a solid support, to mimic a bioactive species immobilized on a biomaterial surface. The carbohydrate recognition processes that take place in the ECM have a pivotal role in promoting cell adhesion and differentiation and, thus, tissue regeneration Hence, we prepared a pseudo-receptor, that mimics the lectin binding site and we characterize its interaction with a panel of different monosaccharides. The results obtained support the theoretical model according to which lectins bind carbohydrates exploiting the CH-π interactions occurring in their active site. Moreover the NMR experimental approach here described can be generally applied when the interacting

species do not have the same solubility properties in physiological conditions and, in particular, can be exploited for the analysis and characterization of molecular recognition events occurring at biomaterial surface.

Keywords: biomaterials, NMR Spectroscopy, STD-NMR, molecular recognition, HR-MAS, carbohydrates, lectins

Natural and synthetic biomaterials represent a powerful tool for the repair of damaged or dysfunctional tissues, both in cell-based and acellular therapeutic strategies. The generation of smart biomaterials for tissue engineering requires mimicking natural components of the extracellular matrix (ECM) involved in the regulation of complex morphogenetic processes responsible for tissue formation and regeneration. To this purpose, a biocompatible scaffold has to be decorated with molecules able to promote cell adhesion and proliferation and/or differentiation through molecular recognition processes. With the aim to obtain cell- and tissue-specificity, the functionalities introduced have to be modulated according to a specific biological environment.

Receptor-ligand interaction studies allow collecting fundamental information on the binding events at the base of these biological phemomena. In particular, the acquired knowledge is necessary to identify natural ligands suitable for the biocompatible scaffolds decoration and to verify that, after immobilisation on the biomaterial surface, ligands retain the ability to interact with the specific biological receptor. Many techniques are currently employed to gain this information, including theoretical and experimental protocols.

In the last few years, NMR has emerged as one of the essential techniques to perform molecular recognition studies, at different levels of complexity,

and employing different systems, from pure samples to in-cell approaches. The characterization of interaction events that take place at biomaterial surface requires the study of non-homogenous samples.

From the NMR perspective, these materials might not be soluble or homogeneous enough for performing solution NMR studies in physiological conditions (aqueous media). For this reason, we feel that the development and application of new NMR tools for the study of receptor-ligand interactions in non-homogenous media has strategic importance.

In particular, we will herein focus on exploring the potential of the use of HR-MAS (High Resolution Magic Angle Spinning) techniques.

In a HR-MAS experiment, the sample is spun at a high speed (typically 3–5 kHz) around an axis oriented at the "magic angle", that is an angle of 54.7° respect to the direction of the magnetic field. This magic angle spinning allows to average the anisotropic interactions, resulting in obtaining high resolution NMR spectra, characterized by a substantial line narrowing that is comparable with liquid state NMR spectra.[1, 2]

Among the NMR methods based on the ligand observation, one of the most robust and versatile experiments is the Saturation Transfer Difference (STD)-NMR.[3] In the last decade, STD-NMR spectroscopy has been extensively exploited to study receptor-ligand interactions.[4-13] To the best of our knowledge, only few examples of STDs have been performed by employing HR-MAS techniques with immobilized ligands or receptors. Klein et al. described the use of HR-MAS STD-NMR to screen the affinity of a mixture of oligosaccharides to a model natural lectin, wheat germ agglutinin (WGA), immobilized to controlled pore glass (CPG).[14] Friebolin et al. employed this technique to characterize the binding of a mixture of eight steroids with an HPLC stationary phase functionalized with cholesterol.[15] Soubias et al., exploited STD under MAS condition to probe

specific lipid-protein interactions;[16] and finally, very recently, we have demonstrated that the combination of the STD experiment with the use of a HR-MAS probe improved the versatility of the STD experiment, expanding its use to samples containing living cells deriving from solid tissues.[17]

Here, we have decided to generate an artificial synthetic model-system, which can be of general utility to verify and extend the application of this approach. The system gathers conceptual interest at both the basic and oriented perspectives, using the key feature of protein-carbohydrate recognition (CH-π interactions), reinforced through multivalent interactions, and also opens new developments in the application of NMR methodologies to the design, generation and characterization of biomaterials.

The basic receptors for sugars in Nature are lectins, which show a high specificity for particular sugar moieties and are involved in a lot of recognition processes in the ECM. They bind to carbohydrates, which are usually forming part of glycoconjugates, as glycoproteins or glycolipids. They achieve their affinity and specificity by the synergic combination of hydrogen bonds and non-polar interactions,[18] especially van der Waals forces and sugar-aromatic (CH–π interactions involving the side chains of aromatic aminoacids and the non polar faces of the carbohydrates). In fact, clusters of non-polar CH bonds at the sugar pyranose rings may form a hydrophobic surface able to generate a stabilizing stacking interaction with the aromatic residues at the protein biding site.[19,20] Interestingly, it has been demonstrated, both experimentally and theoretically, that certain simple sugars, with three CH bonds pointing towards the same spatial area, may interact with the aromatic surface of simple aromatic rings, such as benzene or phenol, even without the necessity of being accommodated in a protein binding site.[21-25]

We herein present our studies on the interaction between a panel of methyl-pyranosides, characterized by displaying different extended hydrophobic surfaces, (Fig. 1) and an “artificial” lectin mimic, just formed by a Sepharose-based resin to which L-tryptophan (Trp) residues have been attached.

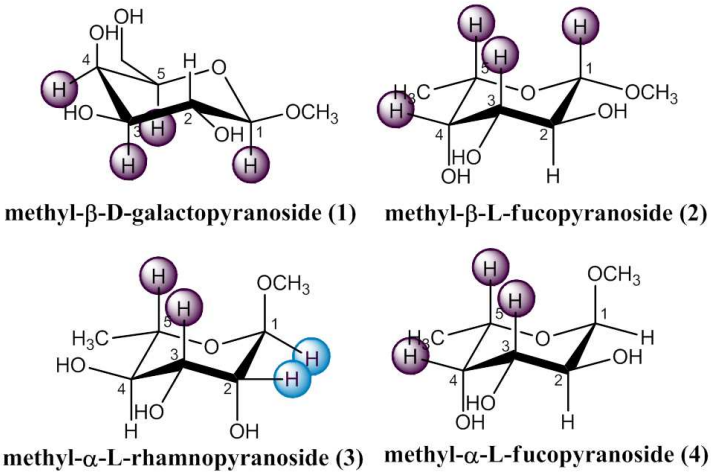


Fig. 1 The hydrophobic faces of the methyl-pyranosides analyzed in this study.

The interactions present in this non-homogeneous system have been monitored by using STD HR-MAS NMR to expand the applicability of the method.

2. Materials and Methods

2.1 Synthesis of pyranosides

The pyranosides were obtained in a near-quantitative yield by a standard Fischer glycosylation from the corresponding monosaccharide. In particular, to a stirring solution of the monosaccharide in dry methanol (MeOH), 0.2 equivalents of acetylchloride were added, with the consequent in situ formation of HCl. The reaction mixture was stirred at room temperature

until the complete conversion of the starting material. The reaction mixture was neutralized with Dowex resin 1x8-100 and then filtered and concentrated under reduced pressure.

2.2 NMR spectroscopy

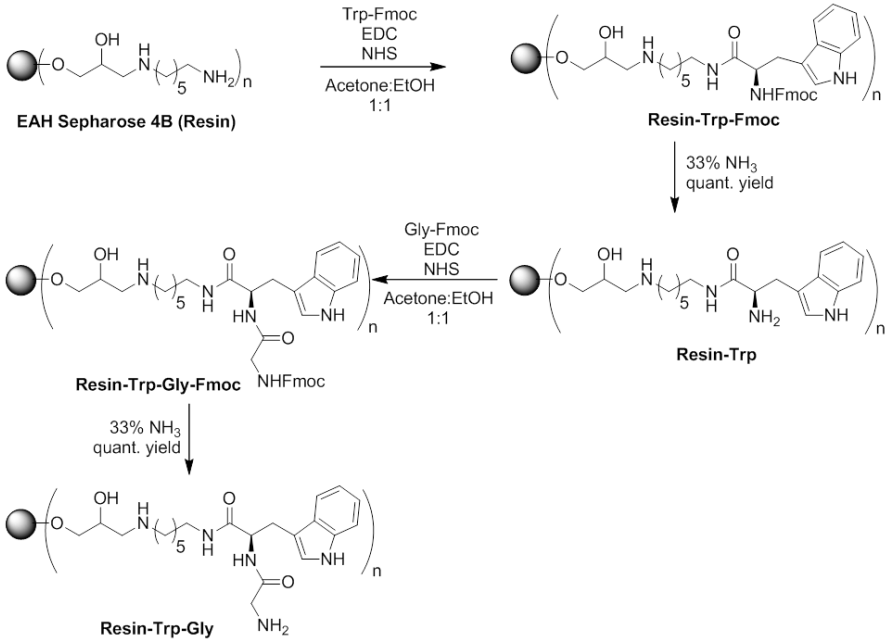
NMR spectra were acquired using a Bruker Advance III 600 MHz NMR spectrometer, equipped with a 4mm TXI HR-MAS probe. Ligand resonances were assigned by acquiring 1- and 2-D standard spectra (^{1}H, ^{1}H-^{1}H COSY, ^{1}H-^{1}H NOESY, ^{1}H-^{13}C HSQC) with the pulse sequences provided by the manufacturer. STD spectra were acquired by applying the standard pulse sequence, with an on-resonance frequency of 7.58, 7.44, 7.14 or 0.0 ppm and an off-resonance frequency of 30.0 ppm. A train of 40 Gaussian-shaped pulses of 50 ms each was employed, separated by a 1 ms delay, with a total saturation time of the protein envelope of 2 s. All samples were prepared by putting 1mg of Resin, Resin-Trp or Resin-Trp-Gly into a disposable KelF insert, adding a solution of the ligand (or of a ligand mixture) in D_2O to a final concentration of 10 mM and then placing it into a zirconia 4 mm rotor. Each sample, with a total volume of 40µl, was maintained at a temperature of 298K during the NMR spectrum acquisition.

3. Results and discussion

Trp immobilization on the solid support allowed the small molecule to acquire relaxation properties typical of a macromolecule, thus mimicking the Trp residues present in a protein structure.

In particular, the Fmoc-Trp moieties were linked to EAH Sepharose 4B® resin (GE Healthcare) presenting a 10-carbon atoms aminated spacer (hereinafter referred to as Resin) by a standard solid-phase peptide synthesis (SPPS) procedure (Scheme 1).

The Fmoc-Trp acidic function was activated by treatment with 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide (EDC) and N-hydroxysuccinimide (NHS) in acetone: ethanol 1:1. The coupling reaction was performed by adding the mixture to a syringe containing the 2 mL of Resin and stirring at r.t. over night. Resin-Trp-Fmoc was then stirred in a 33% NH3 solution for 2 h, to remove Fmoc protective group, affording Resin-Trp. The same procedure was applied to prepare an alternative receptor by linking a Gly residue to Resin-Trp, to have a receptor model in which both the Trp carboxylic and amino groups are involved in a peptidic bond. The resins were lyophilized and then stored at 4°C. The 2,4,6-tri-nitrobenzenesulfonic acid (TNBS) test [26] was used in order to monitor the overall process. Then, the ability of the different methyl-pyranosides (Fig. 1) to bind Trp immobilized on Resin was investigated by STD-NMR experiments.



Scheme 1 Functionalization of EAH Sepharose 4B® resin by conventional SPPS.

The samples were prepared by setting ca. 1 mg of resin (Resin, Resin-Trp, or Resin-Trp-Gly) into a 4mm MAS rotor and subsequently filling it with a solution of monosaccharide in D_2O to a final concentration of 10mM (the estimated Trp : monosaccharide ratio was 1:15). The samples were spinned at 4 kHz and the ^1H and STD spectra were acquired on a Bruker 600 MHz spectrometer.

Methyl-β-D-galactopyranoside (**1**), which presents an extended non polar face (Fig. 1), binds both Resin-Trp and Resin-Trp-Gly, as shown by the STD spectra reported in Fig. 2.

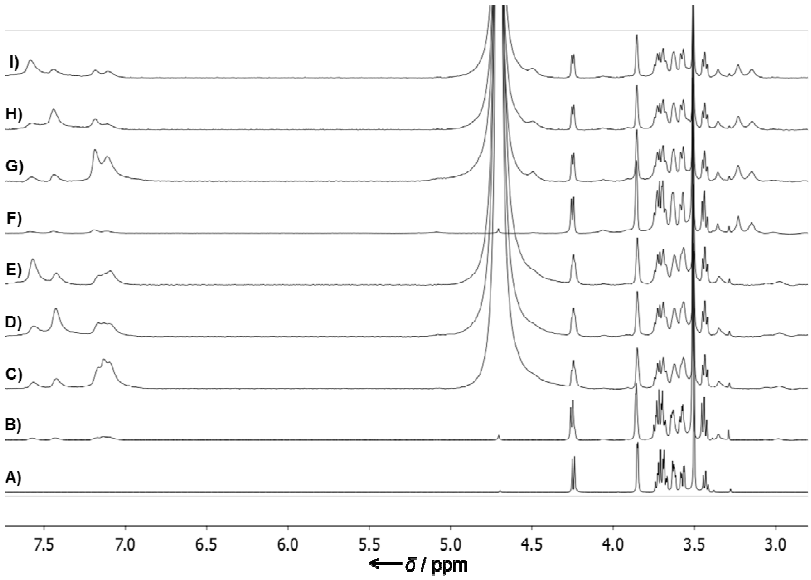
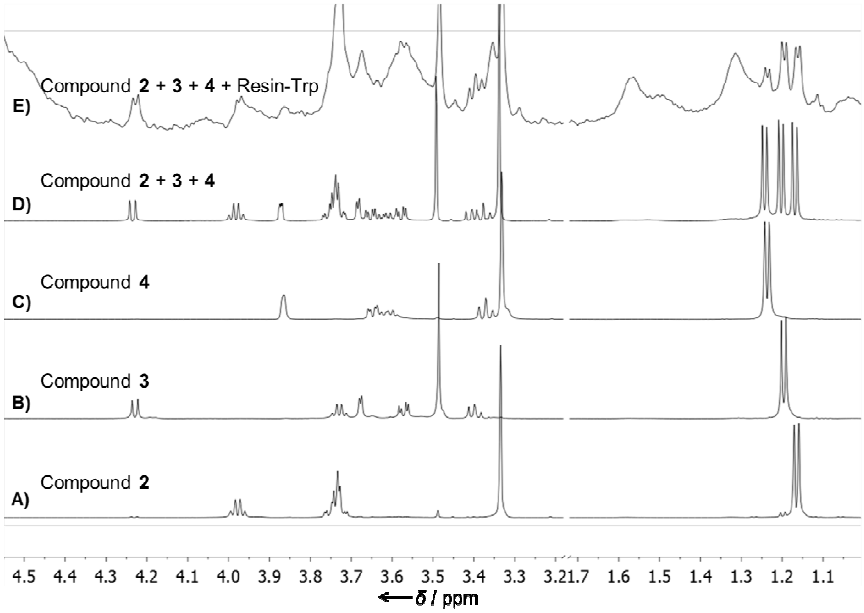


Fig. 2 A) ^1H spectrum of methyl-β-galactopyranoside (**1**) in D_2O; B) ^1H spectrum of the Resin-Trp/methyl-β-galactopyranoside (**1**) mixture in D_2O; C-E) 1D-STD spectra of the mixture in D_2O. Selective saturation time 2s, recorded with different selective saturation frequencies (C, 7.14 ppm; D, 7.44 ppm; E, 7.58 ppm); F) ^1H spectrum of the Resin-Trp-Gly/methyl-β-galactopyranoside (**1**) mixture in D_2O; G-I) 1D-STD spectra of the mixture in D_2O. Selective saturation time 2s, recorded with different selective saturation frequencies (G, 7.14 ppm: H, 7.44 ppm; I, 7.58 ppm).

Indeed, when the Trp resonances were selectively irradiated, the methyl-β-galactopyranoside (**1**) signals were clearly present in the STD spectra. Thus,

a magnetization transfer process from the Trp-functionalized resin to the ligand is taking place, indicating the existence of a sugar/resin binding process. It is not unfair to assume that the binding is due to interactions between the Trp residues and the sugar moiety. Indeed, when the same experiment was performed in the presence of the non functionalized resin (Resin), the selective irradiation of the resin aliphatic resonances did not promote the presence of any sugar signals in the corresponding STD



spectrum (Fig. 3).

Fig. 3 A) ^{1}H spectrum of the Resin-Trp/methyl-α-L-fucopyranoside (**2**) mixture in D_2O; B) ^{1}H spectrum of the Resin-Trp/methyl-β-L-fucopyranoside (**3**) mixture in D_2O; C) ^{1}H spectrum of the Resin-Trp/methyl-α-L-rhamnopyranoside (**4**) mixture in D_2O; D) ^{1}H spectrum of the Resin-Trp/methyl-α-L-fucopyranoside (**2**)/methyl-β-L-fucopyranoside (**3**)/methyl-α-L-rhamnopyranoside (**4**) mixture (1:1:1 for the sugars) in D_2O; E) 1D-STD spectrum of the Resin-Trp/methyl-α-L-fucopyranoside (**2**)/methyl-β-L-fucopyranoside (**3**)/methyl-α-L-rhamnopyranoside (**4**) mixture (1:1:1 for the sugars) in D_2O, selective saturation time 2s, selective saturation frequency 7.14ppm.

Methyl-α-L-fucopyranoside (**2**) and methyl-β-L-fucopyranoside (**3**) present a non polar surface similar to that of methyl-β-D-galactopyranoside (**1**), (they

are 6-deoxy galactopyranosides), while methyl-α-L-rhamnopyranoside (**4**) displays a different topology (Fig. 1). Thus, we repeated the same experimental approach employing these saccharides. They specifically interacted with the Trp residue, as demonstrated by the appearance of their signals in the STD experiments recorded in the presence of both Resin-Trp and Resin-Trp-Gly, while there was no evidence of binding with Resin alone (Fig. 3).

Nevertheless, on the basis of their CH bond spatial disposition, some differences in their binding affinity should be expected. To test this hypothesis, a competitive STD was acquired on a mixture of the three sugars at a 1:1:1 molar ratio. As shown in Fig. 4, the rhamnopyranoside (**4**) clearly interacted significantly more weakly than the two fucosides (**2** and **3**). Indeed, STD signals of the methyl groups of this molecule were one third in intensity of those corresponding to the two fucosides.

STD HR-MAS NMR experiments in heterogeneous phase may be exploited to monitor molecular recognition events that occur on biomaterial surface. The study presented herein mimics the process of pyranoside binding to lectins that takes place in Nature. We have used the CH–π interactions formed between simple sugars and an artificial system synthesised from a resin by decoration with aromatic residues. This example shows that this methodology can be applied even when the interacting species do not show the same physical properties, or when the experimental conditions prevent the application of high resolution liquid NMR techniques.

The success of the experiment takes also advantage of one key concept in molecular recognition: multivalency. In this example, the covalent linkage of many Trp molecules to the resin involves the“multi-exposition” of the interacting species on a solid surface. The “multi-exposition” of the same ligand and/or receptor is a characteristic feature of the interactions taking

place on cell surface and in the ECM and in which carbohydrates and lectins play a pivotal role. Thus, the proposed method appears even more suitable for the study of recognition events involving biomolecules employed for biomaterial decoration. At the same time, this chemical approach exploits the advantages of multivalency, by increasing the local concentration of one of the two “actors” involved in the binding, and allows the study of weak interactions, as those presented here. The possibility of expanding this concept for selective recognition of sugars remains an open question.

Acknowledgements

We gratefully acknowledge Fondazione CARIPLO, project 2008/3175, and NEDD (Regione Lombardia, FONDO PER LA PROMOZIONE DI ACCORDI ISTITUZIONALI, Progetto n. 4779 "Network Enabled Drug Design") for financial support and Eleonora Macchi for her experimental work.

References

[1] S. Aime, E. Bruno, C. Cabella, S. Colombatto, G. Digilio, and V. Mainero, HR-MAS of cells: A “cellular water shift” due to water-protein interactions?, Magnetic Resonance in Medicine, 54, (2005), 1547-1552.

[2] M. Coles, M. Heller and H. Kessler, NMR-based screening technologies, Drug Discovery Today 8 (17) (2003) 803-810.

[3] M. Mayer, B. Meyer, Characterization of ligand binding by saturation transfer difference NMR spectra, Angew. Chem. 1999, 111, 1902–1906; Angew. Chem. Int. Ed. 38 (1999) 1784-1788.

[4] T. Biet, T. Peters, Molecular Recognition of UDP-Gal by β-1,4-Galactosyltransferase T1, Angew. Chem. 113 (2001) 4320–4323; Angew. Chem. Int. Ed., 40 (2001) 4189-4192.

[5] Y. Yuan, D. V. Bleile, X. Wen, D. A. R. Sanders, K. Itoh, H. Liu and B. M. Pinto, Investigation of binding of UDP-Galf and UDP-[3-F]Galf to UDP-galactopyranose mutase by STD-NMR spectroscopy, molecular dynamics, and CORCEMA-ST calculations, J. Am. Chem. Soc., 130 (2008) 3157-3168.

[6] A. Bernardi, D. Arosio, D. Potenza, I. Sanchez-Medina, S. Mari, J. F. Cañada, J. Jimenez-Barbero, Intramolecular Carbohydrate – Aromatic Interactions and Intermolecular van der Waals Interactions Enhance the Molecular Recognition Ability of GM1 Glycomimetics for Cholera Toxin, Chem. Eur. J. 10 (2004) 4395-4406.

[7] A. Canales, R. Matesanz, N. M. Gardner, J. M. Andreu, I. Paterson, J. F. Díaz, J. Jiménez-Barbero, The Bound Conformation of Microtubule-Stabilizing Agents: NMR Insights into the Bioactive 3D Structure of Discodermolide and Dictyostatin, Chem. Eur. J. 14 (2008) 7557–7569.

[8] R. Caraballo, H. Dong, J. P. Ribeiro, J. Jiménez-Barbero, O. Ramström, Direct STD NMR Identification of β-Galactosidase Inhibitors from a Virtual Dynamic Hemithioacetal System, Angew. Chem. 122 (2010) 599-603; Angew. Chem. Int. Ed. 49 (2010) 589-593.

[9] T. Haselhorst, T. Fiebig, J. C. Dyason, F. E. Fleming, H. Blanchard, B. S. Coulson, M. von Itzstein, Recognition of the GM3 Ganglioside Glycan by Rhesus Rotavirus Particles, Angew. Chem. 50 (2011) 1055-1058; Angew. Chem. Int. Ed. 5 (2011) 1055–1058.

[10] C. Airoldi, S. Sommaruga, S. Merlo, P. Sperandeo, L. Cipolla, A. Polissi, F. Nicotra, Targeting Bacterial Membranes: NMR Spectroscopy Characterization of Substrate Recognition and Binding Requirements of D-Arabinose-5-Phosphate Isomerase, Chem. Eur. J. 16 (6) (2010) 1897-1902.

[11] C. Airoldi, F. Cardona, E. Sironi, L. Colombo, M. Salmona, A. Silva, F. Nicotra, B. La Ferla, cis-Glyco-fused benzopyran compounds as new amyloid-beta peptide ligands ChemComm, 47 (2011) 10266–10268.

[12] C. Airoldi, L. Colombo, C. Manzoni, E. Sironi, A. Natalello, S. M. Doglia, G. Forloni, F. Tagliavini, E. Del Favero, L. Cantù, F. Nicotra, M. Salmona, Tetracycline prevents Aβ oligomer toxicity through an atypical supramolecular interaction, Org. Biomol. Chem. 9 (2011) 463-472.

[13] B. Claasen, M. Axmann, R. Meinecke, B. Meyer, Direct observation of ligand binding to membrane proteins in living cells by STDD NMR, J. Am. Chem. Soc. 127 (2005) 916-919.

[14] J. Klein, R.Meinecke, M. Mayer, B. Meyer, Detecting binding affinity to immobilized receptor proteins in compound libraries by HR-MAS STD NMR, J. Am. Chem. Soc., 121 (1999) 5336-5337.

[15] V. Friebolin, M. P. Bayer, M. T. Matyska, J. J. Pesek, K. Albert, 1H HR/MAS NMR in the suspended state: Molecular recognition processes in liquid chromatography between steroids and a silica hydride-based cholesterol phase J. Sep. Sci. 32 (2009) 1722-1728.

[16] O. Soubias, K. Gawrisch, Probing specific lipid-protein interaction by saturation transfer difference NMR spectroscopy, J. Am. Chem. Soc. 127 (2005) 13110-13111.

[17] C. Airoldi, S. Giovannardi, B. La Ferla, J. Jiménez-Barbero, F. Nicotra, Saturation Transfer Difference NMR Experiments of Membrane Proteins in Living

Cells under HR-MAS Conditions: The Interaction of the SGLT1 Co-transporter with Its Ligands, Chem. Eur. J. 17 (2011) 13395-13399.

[18] R. U. Lemieux, How Water Provides the Impetus for Molecular Recognition in Aqueous Solution, Accounts of Chemical Research 29 (1996) 373-380.

[19] X. Robert, R. Haser, T.E. Gottschalk, F. Ratajczak, H. Driguez, B. Svensson, N. Aghajari, The structure of barley alpha-amylase isozyme 1 reveals a novel role of domain C in substrate recognition and binding: a pair of sugar tongs, Structure 11 (2003) 973.

[20] R. Mahesh Kumar, M. Elango, and V. Subramanian, Carbohydrate-Aromatic Interactions: The Role of Curvature on XH···π Interactions, J. Phys. Chem. A 114 (2010) 4313.

[21] M.S. Sujatha, Y. U. Sasidhar, P.V. Balaji, Energetics of galactose- and glucose-aromatic amino acid interactions: implications for binding in galactose-specific proteins, Protein Sci. 13 (2004) 2502.

[22] M.C. Fernandez-Alonso, F. J. Canada, J. Jimenez-Barbero, G. Cuevas, Molecular Recognition of Saccharides by Proteins. Insights on the Origin of the Carbohydrate−Aromatic Interactions, J. Am. Chem. Soc. 127 (2005) 7379.

[23] J. Screen, E.C. Stanca-Kaposta, D.P. Gamblin, B. Liu, N.A. Macleod, L.C. Snoek, B.G. Davis, J.P. Simons, Angew. Chem. Int. Ed. 46 (2007) 3644–3648.

[24] Sophie Vandenbussche, Dolores Diaz, Maria Carmen Fernandez-Alonso, Weidong Pan, Stephane P. Vincent, Gabriel Cuevas, Francisco Javier Canada, Jesus Jimenez-Barbero, and Kristin Bartik, Aromatic–Carbohydrate Interactions: An NMR and Computational Study of Model Systems, Chem. Eur. J. 14 (2008) 7570.

[25] S. Kozmon, R. Matuška, V. Spiwok, J. Koča, Three-Dimensional Potential Energy Surface of Selected Carbohydrates' CH/pi Dispersion Interactions Calculated by High-Level Quantum Mechanical Methods, Chem. Eur. J. 17 (2011) 5680–5690.

[26] W.S Hancock, J.E. Battersby, A new micro-test for the detection of incomplete coupling reactions in solid-phase peptide synthesis using 2,4,6-trinitrobenzene-sulphonic acid, Anal. Biochem. 71 (1976) 260-264.

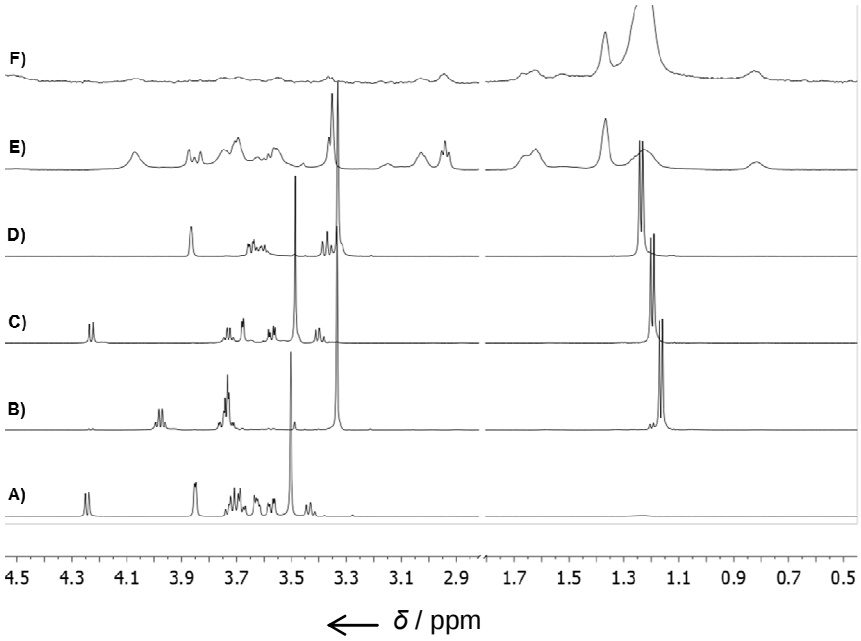


Figure S1. A) ^{1}H spectrum of methyl-β-D-galactopyranoside in D_2O; B) ^{1}H spectrum of methyl-α-L-fucopyranoside in D_2O; C) ^{1}H spectrum of methyl-β-L-fucopyranoside in D_2O; D) ^{1}H spectrum of methyl-α-L-rhamnopyranoside mixture in D_2O; E) ^{1}H spectrum of Resin in D_2O (5x); F) 1D-STD spectrum of the Resin/methyl-β-D-galactopyranoside/methyl-α-L-fucopyranoside/methyl-β-L-fucopyranoside/methyl-α-L-rhamnopyranoside mixture in D_2O, selective saturation time 2s, selective saturation frequency 0.0 ppm (10x).

Concluding Remarks

General Conclusions

The NMR techniques useful for the investigation of different recognition processes have been successfully applied. In particular

- a versatile protocol based on STD experiments for the study of KdsD binding requirements and screening of putative inhibitors, towards the rational design and optimization of novel antibacterial agents,
- a reliable method for the screening of PrP binding compounds, towards the development of effective diagnostic or putative therapeutic compounds,
- a NMR approach for the study of interaction events occurring in non-homogeneous conditions,

have been provided.

Moreover some more insights on the binding of OGP to α2M have been presented.

Riassunto

La stragrande maggioranza dei processi biologici si basa su interazioni che avvengono tra macromolecole o tra piccoli ligandi e proteine. Di conseguenza, la caratterizzazione degli eventi di riconoscimento molecolare è fondamentale per avere una maggiore comprensione dei meccanismi alla base di fenomeni di interesse biologico e biomedico; tali conoscenze infatti possono poi essere sfruttate per la progettazione razionale di nuovi strumenti terapeutici e/o diagnostici. In questo contesto la spettroscopia NMR (Risonanza Magnetica Nucleare) è uno strumento ampiamente utilizzato e versatile, in quanto fornisce informazioni a livello atomico, utili per lo studio degli aspetti strutturali e dinamici che coinvolgono le biomolecole e che riguardano le loro interazioni in soluzione.

In questa tesi sono state studiate, tramite tecniche NMR, diverse coppie ligandi-proteina di rilevanza biologica e biomedica.

Il primo sistema studiato è stato rappresentato dall'enzima arabinosio-5-fosfato isomerasi (API o KdsD) e la sua interazione con i substrati naturali e loro derivati. È il primo enzima coinvolto nella biosintesi dell'acido 2-cheto-3-desossi-D-manno-ottulosonico (KDO), e catalizza l'isomerizzazione reversibile da D-ribulosio 5-fosfato (Ru5P) a D-arabinosio 5-fosfato (A5P). Poiché l'inibizione della biosintesi del KDO porta all'arresto della crescita cellulare e poiché non esiste una controparte umana di KdsD, quest'ultima rappresenta un bersaglio ideale per lo sviluppo di nuovi composti antimicrobici.

Innanzitutto è stato messo a punto un metodo NMR per caratterizzare le proprietà catalitiche e di legame dei substrati degli enzimi KdsD di *E. coli* e *P. aeruginosa* (Ec-KdsD e Pa-KdsD rispettivamente). Ec-KdsD e Pa-KdsD hanno un'identità di sequenza del 57% e similarità di sequenza del 75% e, attraverso gli studi di interazione che abbiamo effettuato su un piccolo gruppo di analoghi dei substrati, abbiamo verificato che Ec-KdsD e Pa-KdsD hanno requisiti di legame del tutto analoghi. In particolare, abbiamo dimostrato che (1) la stereochimica corretta in posizione 2 non è necessaria per il riconoscimento, ma è necessaria per la catalisi enzimatica; (2) la posizione 3 è coinvolta nel riconoscimento e la sua stereochimica corretta è fondamentale per il legame dell'enzima; (3) il gruppo ossidrile in posizione 4 non è necessario né per l'interazione con KdsD, né perché l'isomerizzazione possa avvenire (4) gruppo fosfato. Inoltre, attraverso la generazione di

alcuni mutanti di Pa-KdsD, sono stati identificati dei residui amminoacidici chiave per l'attività della proteina: dati NMR e biochimici ci hanno portato a ipotizzare un ruolo strutturale per l'istidina 85 di Pa-KdsD e una funzione catalitica per la lisina 56 e l'istidina 190. Al fine di meglio caratterizzare Pa-KdsD e approfondire i suoi requisiti di legame, necessari per la progettazione di nuovi inibitori, abbiamo intrapreso uno studio di riconoscimento molecolare di due dei più potenti inibitori trovati fino ad oggi (composti **1** e **2**), insieme ad alcuni altri derivati dell'arabinosio 5-fosfato (Fig.1).

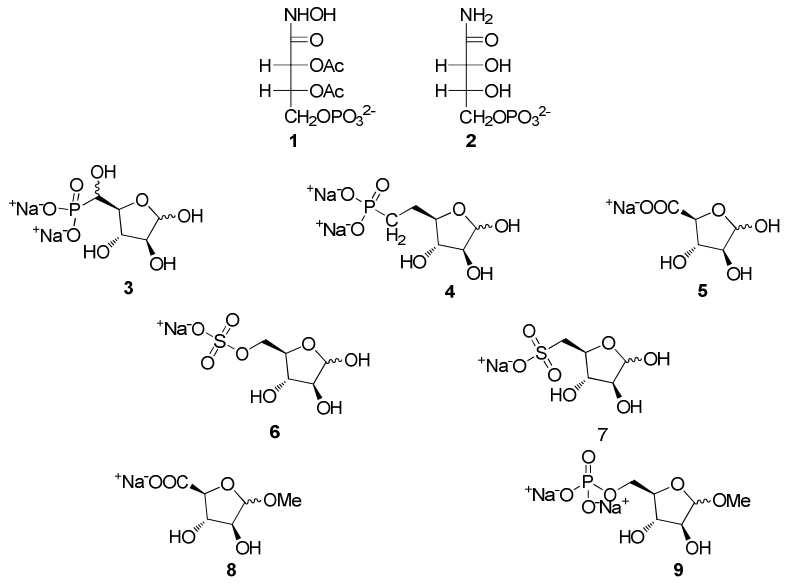


Figura 1 . Struttura dei composti testati per il loro legame a Pa-KdsD.

Sulla base dei risultati ottenuti, si può affermare che i due inibitori noti effettivamente interagiscono con KdsD e probabilmente il composto **1** ha una potenza superiore a quella precedentemente riportata; il carbossilato, solfato, solfonato e il gruppo $-CH_2(OH)PO_3^{2-}$ non sono bio-mimetici, mentre il fosfonato e i gruppi O-Metilici in posizione anomerica non impediscono il legame a KdsD, ma la loro presenza ne influenza l'affinità, che risulta essere inferiore a quella dei substrati naturali dell'enzima. I risultati biologici preliminari dimostrano che né composti i **8**, né i composti **9** esercitano un effetto inibitorio, alle concentrazioni testate, sulla crescita dei ceppi selvatici di *P. aeruginosa* e di *E. coli*, ma tali composti hanno attività

batteriostatica se testati sul mutante di *E. coli* AS19, caratterizzato da una parete cellulare difettosa. E' in corso la caratterizzazione biologica degli altri compositi sintetizzati, e alcuni altri potenziali inibitori sono in preparazione.
La seconda tematica riguarda la proteina prionica PrP. Le malattie da prioni sono patologie letali, caratterizzate da un accumulo cerebrale di forme conformazionalmente modificate della proteina prionica cellulare (PrPC), note come PrPSc. Infatti, mentre PrPC è prevalentemente ricca di α-eliche, PrPSc è prevalentemente costituita da strutture a foglietto-β. Studi precedenti hanno dimostrato che un peptide sintetico corrispondente ai residui 106-126 della variante umana (PrP106-126) presenta alcune delle proprietà chimico-fisiche e patogeniche di PrPSc; inoltre, tale regione è necessaria per la conversione di PrPC in PrPSc. Sebbene ad oggi non esistano ancora agenti diagnostici e terapie efficaci per questa patologia, sono note diverse piccole molecole in grado di legare questi peptidi e inibire la loro aggregazione *in vitro*; la maggior parte di esse sono composti naturali recanti porzioni aromatiche. In particolare è stato riportato che le tetracicline sono in grado di legare il PrP106-126, di inibire l'aggregazione di PrP in fibrille amiloidi, e di prevenire la morte neuronale. Tuttavia tali molecole sono caratterizzate da basse solubilità e stabilità chimica in condizioni fisiologiche e mostrano attività antibiotiche indesiderate in questo contesto. Cercando di superare queste limitazioni, sono stati recentemente sintetizzati nel nostro laboratorio alcuni composti triciclici stabili e solubili, che si sono già dimostrati essere ligandi dei peptidi Aβ, responsabili della malattia neurodegenerativa nota come morbo di Alzheimer. Spesso molecole attive contro un processo amiloidogenico, sono anche attive in altre patologie appartenenti a questa categoria; per questo motivo si è pensato di testare tali composti *via* NMR per la loro capacità di interagire con oligomeri del PrP. Queste tricicline sono caratterizzate da uno *scaffold* glicidico, fuso a due anelli a sei termini, ed è stato dimostrato che, pur mantenendo l'aromaticità necessaria, sono prive dell'attività antibiotica indesiderata delle tetracicline e, inoltre, sono risultate essere ligandi di aggregati del PrP. Tramite l'acquisizione di spettri STD-NMR (*Saturation Transfer Difference*) si è scoperto che minore è la polarità dei sostituenti sull'anello aromatico, maggiore è l'affinità per tali aggregati e che il residuo di zucchero è meno coinvolto nel processo di interazione e può quindi essere sfruttato per ulteriori funzionalizzazioni, volte a

migliorarne la capacità di legame e/o a introdurre nuove proprietà. A tal proposito, è stato preparato un derivato in cui la cumarina è stata legata covalentemente allo zucchero, attraverso una corta catena di polietilenglicole. La cumarina è stata scelta come sonda fluorescente che può essere potenzialmente utilizzata per evidenziare la presenza di fibrille amiloidi *in vivo*. Dopo aver verificato che anche questo derivato mantiene la capacità di legarsi efficacemente al PrP, è stato dimostrato che, grazie alle sue proprietà chimico-fisiche, è anche in grado di permeare attraverso un modello della barriera ematoencefalica, probabilmente sfruttando un meccanismo di diffusione. Questo risultato rappresenta un importante traguardo nello sviluppo di molecole efficaci per la diagnosi precoce di queste gravi patologie. Tutti gli altri test biologici sono in corso.

Il terzo argomento affrontato riguarda lo sviluppo di metodi NMR funzionali a studi per la generazione e caratterizzazione di biomateriali. I biomateriali naturali e sintetici rappresentano un potente strumento per riparare i tessuti danneggiati e la loro generazione per l'ingegneria tissutale richiede di “mimare” le componenti naturali presenti nella matrice extracellulare (ECM), coinvolte nella regolazione dei processi morfogenetici e nella formazione e rigenerazione di un nuovo tessuto. A questo scopo, un’opportuna matrice biocompatibile deve essere decorata con molecole in grado di promuovere l'adesione, la proliferazione e/o il differenziamento cellulare attraverso processi di riconoscimento molecolare. Di conseguenza è necessario identificare ligandi naturali adatti e di verificare che, dopo l'immobilizzazione sulla superficie del biomateriale, tali ligandi mantengano la capacità di interagire con il loro specifico recettore. Inizialmente abbiamo identificato il peptide OGP (*Osteogenic Growth Peptide*) come molecola interessante da studiare: è un peptide naturale di quattordici amino acidi altamente conservato, che ha suscitato notevole interesse come agente anabolico, in quanto svolge effetti regolatori sulle cellule ossee e sul midollo osseo e mostra effetti mitogenici su fibroblasti. È fisiologicamente presente nella circolazione sanguigna, principalmente nella forma di complesso con proteine OGPB (OGP-binding protein). In particolare, l’α-2-macroglobulina (α2M) è una OGPBP plasmatica di 720 kDa, che sembra essere un importante modulatore dell'azione dell’OGP: protegge l’OGP dalla degradazione proteolitica e fornisce un meccanismo per il controllo della sua disponibilità per il legame al suo putativo recettore

bersaglio. Tale recettore non è mai stato identificato e non esistono ancora dati strutturali sulla conformazione del peptide in condizioni fisiologiche e in presenza di α2M. Considerando il ruolo chiave di α2M nella regolazione dell'attività proliferativa dell'OGP, abbiamo applicato tecniche NMR ad alta risoluzione al fine di indagare a livello molecolare la formazione del complesso OGP-α2M. I dati NMR acquisiti mostrano che le regioni dell'OGP maggiormente coinvolte nel legame ad α2M sono principalmente i gruppi metile di Leucina 2 e 9, Treonina 8 e Alanina 1 e le porzioni aromatiche di Fenilalanina 12 e Tirosina 10. Inoltre è stato verificato che l'OGP(1-9) compete con l'OGP per l'interazione con α2M, mentre il frammento OGP(10-14), che risulta essere la minima sequenza derivata dall'OGP che mantiene l'attività biologica, lega α2M in un sito diverso, in quanto non compete con l'OGP per il legame alla OGPB. Abbiamo cercato di indagare la conformazione dell'OGP attraverso l'acquisizione ed analisi di spettri 2D-NOESY e tr-NOESY, ma purtroppo i *constraints* ricavati non sono stati sufficienti a generare una struttura diversa dalla conformazione *random coil* osservata per il peptide in soluzione. Quindi, sebbene la presenza di α2M in qualche modo stabilizzi il peptide, probabilmente le distanze intramolecolari sono superiori a 5Å e quindi non sono osservabili in questo tipo di esperimenti. Tale ipotesi è estremamente verosimile alla luce delle dimensioni relative delle due molecole. Infatti il piccolo peptide (1523 Da) potrebbe accomodarsi nel sito di legame di α2M (720 kDa) senza doversi ripiegare in una conformazione sostanzialmente differente da quella del peptide libero in soluzione. Grazie alle sue proprietà osteogeniche, l'OGP sarà utilizzato per la decorazione di biomateriali per ingegneria tissutale e, non appena tali materiali saranno generati, verranno eseguiti studi NMR direttamente sul materiale.

Nel frattempo, al fine di sviluppare esperimenti STD HR-MAS da impiegare per questo tipo di studi, abbiamo sfruttato le caratteristiche delle interazioni proteine-carboidrati per generare un sistema-modello di riconoscimento molecolare per esperimenti NMR in condizioni eterogenee. Le lectine sono proteine coinvolte in molti processi di riconoscimento che avvengono nella ECM. Esse riconoscono e legano selettivamente i carboidrati che di solito fanno parte di complessi glicoconiugati, come glicoproteine o glicolipidi. La loro affinità e specificità derivano dalla combinazione sinergica di legami idrogeno e interazioni non polari, in

particolare forze di van der Waals e interazioni tra i legami CH degli zuccheri e le catene laterali degli amminoacidi aromatici presenti nella tasca di legame. In questo contesto abbiamo esplorato la possibilità di rivelare l'interazione tra un piccolo gruppo di monosaccaridi, caratterizzati da superfici idrofobiche di diversa estensione, e un residuo di L-triptofano (Trp) covalentemente legato ad una resina di Sefarosio. L'immobilizzazione del Trp su un supporto solido consente alla piccola molecola di acquisire le proprietà di rilassamento tipiche di una macromolecola, e di conseguenza tale residuo imita efficacemente un residuo di Trp presente sulla superficie di in una proteina. I risultati ottenuti supportano il modello secondo cui il legame tra lectine e zuccheri si basa sulle interazioni CH-π che si verificano con la catena laterale di residui aromatici presenti nel loro sito attivo. Infatti abbiamo dimostrato sperimentalmente che ad una maggiore estensione della faccia idrofobica dello zucchero corrisponde un'interazione a più alta affinità con il residuo aromatico. Inoltre, tali esperimenti hanno confermato la possibilità di condurre esperimenti di interazione in fase eterogenea attraverso l'utilizzo di sonde HR-MAS, rendendo questo approccio generalmente applicabile quando le specie interagenti non hanno le stesse proprietà di solubilità in condizioni fisiologiche. Inoltre il successo di questo esperimento sfrutta anche un concetto chiave nei processi di riconoscimento molecolare, ossia la multivalenza, in quanto il legame covalente di molte molecole di Trp alla resina comporta la "multi-esposizione" delle specie che interagiscono su una superficie solida. Questa "multi-esposizione" del ligando stesso e/o del recettore oltre a consentire lo studio di interazioni deboli, è una caratteristica delle interazioni che avvengono sulla superficie cellulare e nella ECM e in cui carboidrati e lectine giocano un ruolo fondamentale. Pertanto, il metodo proposto appare particolarmente adatto per lo studio di eventi di riconoscimento che coinvolgono biomolecole impiegate per la decorazione di biomateriali.

Papers, reviews and book chapter

1. Airoldi C., Merlo S., Sironi E., NMR molecular recognition studies for the elucidation of protein and nucleic acid structure and function, e-book chapter, *accepted*
2. Airoldi C., Merlo S., Sironi E., Nicotra F., Jimenez-Barbero J. NMR Protein-Ligand Interaction Studies under Non-Homogeneous Conditions for Biomaterial Generation: A Model For Artificial Lectin-Carbohydrate Recognition, Journal of Materials Science and Engineering B 2 , 12, **2012**, 618-625
3. La Ferla B., Airoldi C., Zona C., Orsato A., Cardona F., Merlo S., Sironi E., D'Orazio G., Nicotra F. Natural glycoconjugates with antitumor activity, *Nat. Prod. Rep.*, 28, 630-648, **2011**
4. Cipolla L., Polissi A., Airoldi C., Gabrielli L., Merlo S., Nicotra F. New Targets for Antibacterial Design: Kdo Biosynthesis and LPS Machinery Transport to the Cell Surface, *Current Medicinal Chemistry*, 18, 6, 830-852, **2011**
5. Airoldi C., Sommaruga S., Merlo S., Sperandeo P., Cipolla L., Polissi A., Nicotra F. Targeting Bacterial Membranes: identification and NMR Characterization of Substrate Recognition and Binding of *P. aeruginosa* D-Arabinose-5P Isomerase, *ChemBioChem*, 12, 719-727, **2011**
6. Airoldi C., Merlo S., Nicotra F. Synthesis of 3-deoxy-D-*threo*pentofuranose 5-phosphate, a substrate of arabinose 5-phosphate isomerase, *Journal of Carbohydrate Chemistry, 29:30–38,* **2010**
7. Airoldi C., Sommaruga S., Merlo S., Sperandeo P., Cipolla L., Polissi A., Nicotra F. Targeting Bacterial Membranes: NMR characterization of substrate recognition and binding requirements of D-arabinose 5-phosphate isomerase, *Chem. Eur. J.*, **2010**, 16, 1897 – 1902

Two other manuscripts are in preparation.

Oral communication

1. S.Merlo, C.Airoldi, S.Sommaruga. P.Sperandeo, L.Gabrielli, L.Cipolla, A.Polissi, F.Nicotra. Targeting bacterial LPS biosynthesis: NMR characterization of D-arabinose 5-phospate isomerase – natural substrate recognition and binding, Chemical Tools for molecular recognition studies: synthesis and NMR characterization of bioactive molecules, Italian-Spanish Joint Workshop CARIPLO Project project 2008-3175, Milano, 22 Aprile **2010**

Communications

1. Airoldi C., Sironi E., Merlo S., Cardona F., La Ferla B., Colombo L., Messa M., Salmona M., Nicotra F., Towards the development of new tools for the diagnosis and therapy of neurodegenerative diseases: NMR characterization of glycofused tricycles as new amyloidogenic peptide ligands, Convegno Nazionale della Divisione di Chimica dei Sistemi Biologici della SCI, Napoli, Italy ,24-25 September **2012** (Poster)
2. Airoldi C., Merlo S., Sironi E, Nicotra F., Jesús Jiménez-Barbero, Saturation Transfer Difference NMR experiments of membrane proteins in living cells under HR-MAS conditions, Convegno Nazionale della Divisione di Chimica dei Sistemi Biologici della SCI, Napoli, Italy ,24-25 September **2012** (oral communication)
3. Airoldi C., Merlo S., Sironi E.,La Ferla B., Jiménez-Barbero J., Nicotra F., Improving saturation transfer difference NMR experiments of membrane proteins in living cells with HR-MAS NMR, XLI NATIONAL CONGRESS ON MAGNETIC RESONANCE, Pisa, 17-19 September **2012** (Poster)
4. Merlo S., Airoldi C., Sironi E., Macchi E., Nicotra F.; Jimenez-Barbero J., Recognition processes in non-homogeneous media by HR-MAS NMR: artificial lectin interaction with carbohydrates, XLI NATIONAL CONGRESS ON MAGNETIC RESONANCE, Pisa, 17-19 September **2012** (Poster)
5. Sironi E., Merlo S., Airoldi C., Cardona F., La Ferla B., Messa M., Colombo L., Salmona M., Nicotra F., NMR characterization of

glycofused tricycles as ligands for amyloidogenic peptides: going towards the development of new tools for the diagnosis and therapy of neurodegenerative disorders, XLI NATIONAL CONGRESS ON MAGNETIC RESONANCE, Pisa, 17-19 September **2012** (Poster)

6. Merlo S., Airoldi C., Sironi E., Macchi E., Nicotra F.; Jimenez-Barbero J. A model for lectin-carbohydrate recognition exploited for NMR interaction studies in heterogeneous systems, International Carbohydrate Symposium, Madrid, 22-26 July 2012
7. L.Gabrielli, Synthesis and biological activity of cyclic d-arabinose 5-phosphate mimetics against gram-negative bacteria. International Carbohydrate Symposium, Madrid, 22-26 July 2012 (oral communication)
8. Gabrielli L., Merlo S., Airoldi C., Sperandeo P., Polissi A., Nicotra F., Holler T. P., Woodard R. W., Cipolla L. NMR studies on substrate specificity of arabinose 5-phosphate isomerase, International Carbohydrate Symposium, Madrid, 22-26 July **2012** (Poster)
9. Cipolla L., Nicotra F., Merlo S., Airoldi C., Gabrielli L., Design of potential inhibitors of kdo byosynthesis as new antibacterial agents, XIII Convegno-Scuola sulla Chimica dei Carboidrati, Pontignano, 24-27 June **2012** (Poster)
10. Gabrielli L., Bini D., Capitoli A., Sgambato A., Scopini J., Merlo S., Airoldi C., Nicotra F., Cipolla L., Design and synthesis of inhibitors against key carbohydrate processing enzymes, XIII Convegno-Scuola sulla Chimica dei Carboidrati, Pontignano, 24-27 June **2012** (Poster)
11. S.Merlo, C.Airoldi, E. Sironi, F. Nicotra, J. Jimenez-Barbero, NMR Characterization of Osteogenic Growth Peptide Interaction with alpha-2-macroglobulin, SMASH – Small molecules NMR conference, 18-21 September **2011**, Chamonix, Francia (Poster)
12. C.Airoldi, S. Merlo, E. Sironi, E. Macchi, F. Nicotra, J. Jimenez-Barbero, NMR tools for protein-ligand interaction studies under non-homogeneous conditions: the example of lectin-carbohydrate recognition, SMASH – Small molecules NMR conference, 18-21 September **2011**, Chamonix, Francia (Poster)
13. Gabrielli L., Airoldi C., Merlo S., Bini D., Cipolla L., Nicotra F. Potential D-arabinose 5-phosphate isomerase (API) inhibitors: Design and synthesis of D-arabinose 5-phosphate mimetics,5th

Glycan Forum, 10-12 Marzo **2011**, Berlino

14. Gabrielli L., Airoldi C., Merlo S., Bini D., Cipolla L., Nicotra F., Synthesis of D-arabinose 5-phosphate mimetics as potential D-arabinose 5-phosphate isomerase (API) inhibitors, IASOC 2010, 25-29 Settembre **2010**, Ischia (Poster)
15. Airoldi C., Merlo S., Sironi E., Macchi E., Nicotra F., Jimenez-Barbero J. Novel NMR applications for characterizing ligand-receptor interactions under non-homogenous conditions, 2010 GERMN-RSEQ meeting, 26-29 Settembre **2010**, Bilbao (Poster)
16. Airoldi C., Sironi E., Merlo S., Macchi E., Nicotra F., Jimenez-Barbero J. New NMR tools for the study of receptor-ligand interactions in non-homogenous media, 14th International Biotechnology Symposium and Exhibition, Biotechnology for the Sustainability of Human Society, 14-18 Settembre **2010**, Rimini (Poster)
17. Airoldi C., Merlo S., Sommaruga S., Sperandeo P., Cipolla L., Polissi A., Nicotra F. Targeting Bacterial Membranes: NMR Characterization of Substrate Recognition and Binding Requirements of D-arabinose 5P Isomerase, a key enzyme in the biosynthesis of LPS, Joint EUROMAR 2010 and 17th ISMAR Conference, 4-9 Luglio 2010, Firenze (Poster)
18. Gabrielli L., Merlo S., Airoldi C., Bini D., Shaikh N., Russo L., Cipolla L., Nicotra F. Synthesis of ligands for Carbohydrate processing enzymes of relevant biological interest, Chemical Tools for molecular recognition studies: synthesis and NMR characterization of bioactive molecules, Italian-Spanish Joint Workshop CARIPLO Project project 2008-3175, Milano, 22 Aprile 2010 (oral communication)
19. Russo L., Shaikh N., Airoldi C., Merlo S., Sironi E., Bini D., Gabrielli L., Cipolla L., Nicotra F. Design of Smart Biomaterials, Chemical Tools for molecular recognition studies: synthesis and NMR characterization of bioactive molecules, Italian-Spanish Joint Workshop CARIPLO Project project 2008-3175, Milano, 22 Aprile 2010 (oral communication)