Ernest Giralt

Design, synthesis and structure of peptides and proteins

Proteins constitute the working machinery and structural support of all organisms. Understanding the molecular basis of a disease, namely where and how the protein network fails, is crucial for developing a therapeutic strategy. The breakthrough concept that proteins function as a contact network rather than as independent agents is not only one of the most important advances in our comprehension of living systems, but it also translates into a new era in drug discovery. The few reported examples of diseases caused by 'impolite' social behaviour of proteins are merely the tip of the iceberg. Therapeutic intervention through molecules designed to selectively modulate the strength and specificity of protein-protein interactions is becoming a reality. Mass spectrometry (MS), nuclear magnetic resonance (NMR) and atomic force microscopy (AFM) are emerging as privileged tools to study the structure of proteins and the complex molecular recognition events that take place in protein-protein interactions, and the interaction of designed ligands – potential new drugs – with protein surfaces. However, the efficient use of these spectroscopic tools is still hampered by numerous difficulties. Regarding MS, the high dynamic range of protein concentrations in living organisms is one of the major challenges in contemporary proteomics. In structure-based drug design, a previous knowledge of the 3D structure of the therapeutic target is mandatory. When this target is a high molecular weight protein, the use of NMR to determine the structure is not possible due to the following two problems associated with this high molecular weight: i) crowding of spectral signals; and ii) fast magnetic relaxation with the subsequent loss of sensitivity. Finally, in amyloid diseases, the study of molecular interactions by AFM or other techniques is strongly hampered by the strong tendency of amyloid proteins to self-assemble.

Using peptidyl aldehydes in activity-based proteomics

Activity-based proteomics (ABP) is a chemical strategy that uses probes that covalently bind the active site of an enzyme. This approach is applied to address protein activity profiling and to discover new therapeutic targets and enzyme inhibitors. Natural substrates and covalent inhibitors of diverse enzymatic classes (*eg*, proteases, glycosidases, and phosphatases) have been used to design several activity-based probes. In most cases, the reactive moiety of these probes consists of electrophilic functional groups such as fluorophosphonates, epoxides, and acyloxymethylketones. Aldehydes are electrophilic functional groups that show inhibitory properties towards several kinds of proteolytic enzymes. Many compounds with an aldehyde moiety have recently been described as covalent reversible inhibitors of serine and cysteine proteases, such as trypsin, thrombin, and cathepsins, among others.

The broad inhibitory spectrum of aldehydes and the possibility that amino acid residues modulate their specificity point to the potential of using peptidyl aldehydes as activity-based probes. We have explored for the first time the potential of peptidyl aldehydes in ABP. For this purpose, we have synthesised various probes and, as a proof of principle, we have used them to specifically label a well-known serine protease in an activity-dependent manner.

Prolyl oligopeptidase (POP; EC 3.4.21.26) is a post-proline serine protease that hydrolyses small proline-containing peptides. POP is involved in the regulation of many bioactive peptides *in vivo*, such as substance P and thyrotropin-releasing hormone, among others, and has been associated with several neuropsychiatric disorders, including schizophrenia and bipolar affective disorder. Although the mechanism of action of this protease remains unknown, several studies have proposed that POP pro-

duces its effect through the metabolism of inositol-1,4,5-triphosphate, a key molecule in the transduction cascade of neuropeptide signalling. In our laboratory, POP was recently cloned from human brain RNA, expressed in *Escherichia coli*, and a homologous model based on the X-ray structure of porcine POP was obtained. To evaluate the use of peptidyl aldehydes in ABP, we synthesised and used three peptidyl aldehyde activity-based probes, Aha-Bpa-Pro-Pro-H (1), Aha-Bpa-Ahx-Pro-Pro-H (2), and Aha-Bpa-Peg-Pro-Pro-H (3) (Aha: hexynoicacid; Ahx: e-aminohexanoic acid; Bpa: benzoylphenylalanine; Peg: 15-amino-4,7,10,13-tetraoxapentadecanoic acid) to label active POP (Figure 1). It is worth mentioning that unlike the phosphonates, sulfonates and other reactive groups traditionally used in activity-based probes, the use of aldehydes facilitates probe synthesis through straightforward and rapid solid-phase peptide synthesis (SPPS) strategies.

Having synthesised the peptidyl aldehyde probes and confirmed their inhibitory properties, we focused on verifying whether these compounds were proper activity-based probes, that is, whether they label POP in an activity-dependent manner and distinguish its active form from heat-denatured samples. To confirm this point, active and heat-denatured POP were incubated with probes 1–3 at 1 and 25 μM for 15 min at rt and crosslinked by UV light irradiation for an additional 60 min at 4 °C. Afterwards, samples were labelled using a trifunctional tag that incorporates both a biotin and a rhodamine moiety49 (TriN3), and were then analysed by SDS–PAGE. These assays showed that peptidyl aldehydes are satisfactory activity-based probes when used at low concentrations. These results were further confirmed with pull-down experiments where the crosslinked probe–protein complex was labelled with the TriN3 tag, extracted from the sample using avidin beads and detected by Western Blot using either an in-house a-POP antibody or streptavidin. This assay was performed with probe 3 (1 μM), and in both cases only the active form of POP was taken from the sample.

Probe 3 was also used for direct mass spectrometry protein identification using an nLC–MS/MS approach. In this case, probe–protein complexes were also labelled with the TriN3 tag, extracted from the sample with avidin, but analysed by reverse-phase chromatography followed by MS/MS identification. MS/MS data were analysed with the SEQUEST software using the IPI database. The maximum false positive rate was set at 1% with PeptideProphet and ProteinProphet. This data analysis led to the identification of POP in active samples but not in heat-denatured controls. These results show that peptidyl aldehydes are adequate activity-based probes not only for specific in-gel

Research Group Members

Group Leader:
Ernest Giralt

Research Associates:
Natàlia Carulla, Teresa Tarragó, Meritxell Teixidó

Postdoctoral Fellows:
Miguel Moreno, Laura Nevola, Óscar Peña, Soledad Royo

PhD Students:
Muriel Arimon, Xavier Arroyo, Andrey Dyachenko, Michael Goldflam, Anna Guimerais, Bernat Guixer, Messim Kichik, Abraham López, Morteza Malakoutikhah, Irene Martín, Laura Mendieta, Roger Prades, Silvia Pujals, Rosa Pujol, Laia Sánchez, Bernat Serra

Research Assistant:
Esther Zurita

Visiting Students:
Vinicius Ilha (Brazil), Nathaly Segovia (Spain)

activity profiling of proteases but also for other ABP applications like protease identification using MS. Unlike other probes that target only one specific type of enzyme at a time (*eg*, serine proteases), peptidyl aldehydes can simultaneously target a variety of proteolytic enzymes.

Therefore, peptidyl aldehydes could provide a valuable tool for identifying unknown proteases by substrate recognition (*eg*, those involved in the activation of neuropeptide precursors), which could eventually lead to the establishment of new therapeutic targets. Moreover, the specificity resulting from the peptide sequence is a crucial feature of peptidyl aldehyde probes and could also be used to successfully monitor the activity of a particular subset of proteins.

A cost-effective labelling strategy for the NMR study of large proteins

NMR spectroscopy is a useful tool for the study of protein structure, protein dynamics and molecular recognition processes, including both protein-protein and protein-ligand interactions. However, the application of NMR experiments to large proteins remains a challenge. Transverse relaxation processes are accelerated as the macromolecule grows and the perdeuteration of proteins with low tumbling rates may be required. Thus, cells must be grown in D_2O, which, in general, reduces protein expression levels and significantly raises the cost of the NMR sample. Moreover, the assignment of spectra is limited by signal overlap. Consequently, simplification of spectra by an appropriate selective labelling scheme is often required. Selective labelling of specific amino acids can be achieved by using auxotrophic cell strains and adding the amino acid with a suitable isotope label to the medium. However, given that the biosynthetic pathways of amino acids are complex, cell growth may be limited when one of these pathways is disrupted, and this leads to lower expression levels.

A less intrusive approach consists of exploiting the cell's metabolic machinery to produce selectively labelled proteins and the particular choice of precursor determines whether a subset or a specific amino acid ends up labelled. Various authors have described such labelling approaches. As an example, when ^{13}C-labelled α-ketobutyrate and α-ketoisovalerate are added to the growth medium, the cell incorporates the ^{13}C-label into valine, leucine and isoleucine side chains. Similarly, addition of [2-^{13}C]- or [4-^{13}C]-labelled indole to the medium allows the labelling of tryptophan residues. Tryptophan, tyrosine and arginine are the most common amino acids in the hot spots of a given protein and are therefore involved in most of the binding energies in protein-protein and protein-ligand interactions. An inexpensive and reliable labelling procedure for the above-mentioned residues that is applicable to large proteins would be extremely useful.

We have set up a cost-effective labelling strategy for the NMR study of large proteins. In this approach, the ^{15}N-label is selectively incorporated into the tryptophan side chains of the protein and the spectrum can be acquired without the need for deuteration. We applied this labelling strategy to POP, a serine protease of 80 kDa (Figure 2).
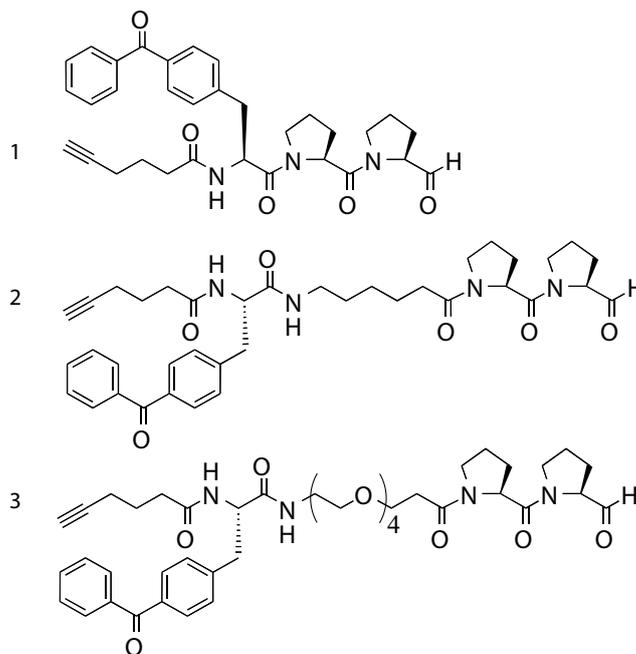

Figure 1. Structure of peptidyl aldehyde probes Aha-Bpa-Pro-Pro-H (1), Aha-Bpa-Ahx-Pro-Pro-H (2), and Aha-Bpa-Peg-Pro-Pro-H (3).
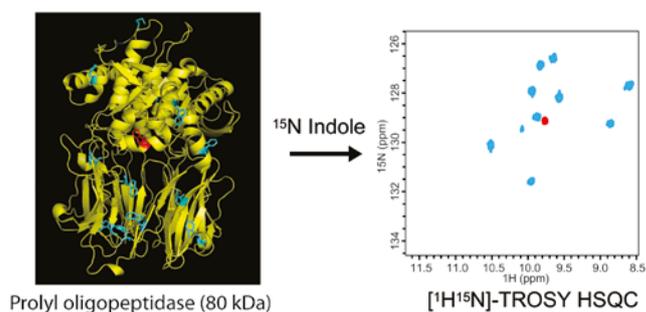

*Figure 2. An ^{15}N-label is selectively incorporated into the tryptophan side chains of a large protein and a simplified [^1H, ^{15}N]-TROSY HSQC spectrum is obtained without the need for deuteration.*

In recent years POP has gained relevance as a target for the treatment of cognitive disturbances. An array of strategies is currently being used to identify POP inhibitors, as these compounds show neuroprotective and cognition-enhancing effects in experimental animals. The X-ray structure of POP from porcine muscle revealed a distinctive two-domain structure: a catalytic domain with an α/β hydrolase fold and an unusual β-propeller domain. The co-structure of POP in the presence of Z-prolyl-prolinal (ZPP), a canonical POP covalent inhibitor, shows that the specificity of the binding between the enzyme and the proline-containing inhibitor is provided by the hydrophobic interaction between POP tryptophan 595 (Trp595) and the ZPP proline ring.
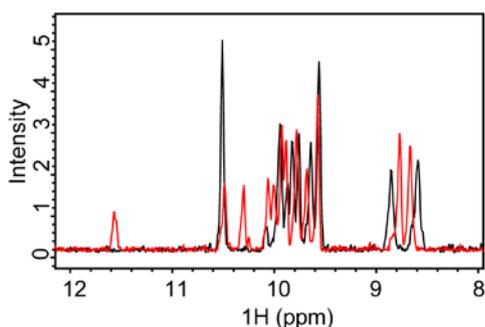
Figure 3. 1H-projection of the [1H,15N]-TROSY HSQC spectra of a selectively Trp[15N-indole]-labelled POP sample. Wild-type POP (black) and POP wild-type with ZPP (red).
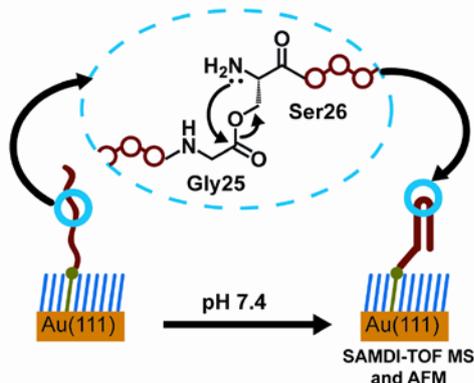
Figure 4. Schematic representation of the intramolecular O-N-acyl migration of Aβ40 on the functionalised gold surface.
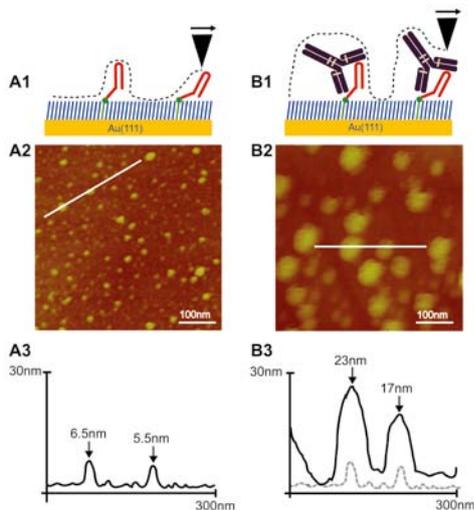
Figure 5. Schematic representation of AFM study of the Aβ40-functionalised gold surface before (A1) and after (B1) incubation; AFM images of the Aβ40-functionalised gold surface before (A2) and after (B2) incubation; surface morphology measurements (AFM) of the Aβ40-functionalised gold surface before (A3) and after (B3) incubation.

The acquisition of a spectrum of a perdeuterated and uniformly 15N-labelled POP sample (U-[2H,15N]-POP) showed that even with perdeuteration and transverse relaxation-optimised spectroscopy (TROSY), signal overlap was still a considerable handicap when dealing with a large protein of 80 kDa. Therefore, a POP sample that was selectively labelled at tryptophan side chains (Trp[15N-indole]-POP) was produced by supplementing the minimal growth medium with 15N-Indole.

The incorporation of the label was checked by MS. The resulting [1H,15N]-TROSY HSQC spectrum recorded with the Trp [15N-Indole]-POP sample showed a satisfactory signal-to-noise ratio (S/N) and eleven of the twelve expected tryptophan signals were present and well-dispersed (Figure 3). It is noteworthy that, despite of the high molecular weight of POP, this result was obtained with a non-perdeuterated protein sample.

Toward the preparation of well-defined amyloid surfaces

Amyloids are a family of self-aggregating proteins related to various central nervous system disorders, including Alzheimer's disease (AD). The two most distinctive brain pathologies of AD are neuritic plaques, which contain mainly insoluble aggregates of the β-amyloid peptide (Aβ), and neurofibrillary tangles of abnormally phosphorylated Tau protein. Cleavage of the β-amyloid precursor protein (APP) by proteases such as β-secretase and γ-secretase produces a predominant product of 40 amino acid residues: β-amyloid peptide (Aβ40). Other Aβ species of various lengths are also produced, including a fibrillogenic peptide of 42 amino acids (Aβ42). The deposition of Aβ40 and Aβ42 in cerebral plaques begins with nucleation of soluble Aβ42 into fibrils, followed by accumulation of soluble Aβ40. In fact, soluble prefibrillar forms of amyloid peptides, including monomers, are suspected to be the main pathogenic factor in AD. Although the mechanism of Aβ toxicity remains unclear, we believe that the study of how Aβ peptides interact with proteins may lead to a better understanding of AD pathogenesis.

We have recently fabricated well-defined, functionalised, monomeric β-amyloid peptide surfaces for the study of protein-protein interactions. We first prepared a non-aggregating analogue of the β-amyloid peptide and then attached it to a gold surface covered with a self-assembled monolayer (SAM) of alkanethiols. After attachment, the native form of the β-amyloid peptide (Aβ40) was obtained by surface-level intramolecular O-N migration (Figure 4). The surface was characterised by atomic force microscopy (AFM) and self-assembled monolayer for matrix-assisted laser desorption/ionization time of-flight mass spectrometry (SAMDI-TOF MS). The interaction between the surface-bound Aβ40 and monoclonal anti-Aβ40 antibody was tracked by AFM and chemiluminescence, which revealed that Aβ40 was attached mainly in its monomeric form and that the protein-protein complex was assembled on the surface (Figure 5). Finally, we used a proteomics approach to demonstrate the specificity of the Aβ40-functionalised surface in surface-binding experiments using serum amyloid P (SAP) and bovine serum albumin (BSA).

Scientific output

Publications

Bastús NG, Sánchez-Tilló E, Pujals S, Farrera C, Kogan MJ, Giralt E, Celada A, Lloberas J and Puntes V. Peptides conjugated to gold nanoparticles induce macrophage activation. *Mol Immunol*, **46**(4), 743-48 (2009)

Bastús NG, Sánchez-Tilló E, Pujals S, Farrera C, López C, Giralt E, Celada A, Lloberas J and Puntes V. Homogeneous conjugation of peptides onto gold nanoparticles enhances macrophage response. *ACS Nano*, **3**(6), 1335-44 (2009)

Boussert S, Diez-Pérez I, Kogan MJ, de Oliveira E and Giralt E. An intramolecular O-N migration reaction on gold surfaces: toward the preparation of well-defined amyloid surfaces. *ACS Nano*, **3**(10), 3091-97 (2009)

Carulla N, Zhou M, Arimon M, Gairí M, Giralt E, Robinson CV and Dobson CM. Experimental characterization of disordered and ordered aggregates populated during the process of amyloid fibril formation. *Proc Natl Acad Sci USA*, **106**(19), 7828-33 (2009)

Comellas G, Kaczmarska Z, Tarragó T, Teixidó M and Giralt E. Exploration of the one-bead one-compound methodology for the design of prolyl oligopeptidase substrates. *PLoS One*, **4**(7), e6222 (2009)

Del Pozo-Rodríguez A, Pujals S, Delgado D, Solinís MA, Gascón AR, Giralt E and Pedraz JL. A proline-rich peptide improves cell transfection of solid lipid nanoparticle-based non-viral vectors. *J Control Release*, **133**(1), 52-59 (2009)

Gaston F, Granados GC, Madurga S, Rabanal F, Lakhdar-Ghazal F, Giralt E and Bahraoui E. Development and characterization of peptidic fusion inhibitors derived from HIV-1 gp41 with partial D-amino acid substitutions. *ChemMedChem*, **4**(4), 570-81 (2009)

Gordo S and Giralt E. Knitting and untying the protein network: modulation of protein ensembles as a therapeutic strategy. *Protein Sci*, **18**(3), 481-93 (2009)

Grillo-Bosch D, Carulla N, Cruz M, Sánchez L, Pujol-Pina R, Madurga S, Rabanal F and Giralt E. Retro-enantio N-methylated peptides as beta-amyloid aggregation inhibitors. *ChemMedChem*, **4**(9), 1488-94 (2009)

Kobayashi S, Nakase I, Kawabata N, Yu HH, Pujals S, Imanishi M, Giralt E and Futaki S. Cytosolic targeting of macromolecules using a pH-dependent fusogenic peptide in combination with cationic liposomes. *Bioconjug Chem*, Epub Apr 23 (2009)

Prieto M, Mayor S, Lloyd-Williams P and Giralt E. Use of the SPhos ligand to suppress racemization in arylpinacolboronate ester Suzuki couplings involving alpha-amino acids. Synthesis of biaryl derivatives of 4-hydroxyphenylglycine, tyrosine, and tryptophan. *J Org Chem*, **74**(23), 9202-05 (2009)

Pujals S, Bastús NG, Pereiro E, López-Iglesias C, Puntes VF, Kogan MJ and Giralt E. Shuttling gold nanoparticles into tumoral cells with an amphipathic proline-rich peptide. *Chembiochem*, **10**(6), 1025-31 (2009)

Sabidó E, Tarragó T and Giralt E. Using peptidyl aldehydes in activity-based proteomics. *Bioorg Med Chem Lett*, **19**(14), 3752-55 (2009)

Sabidó E, Tarragó T, Niessen S, Cravatt BF and Giralt E. Activity-based probes for monitoring postproline protease activity. *Chembiochem*, **10**(14), 2361-66 (2009)

Smrcka AV, Kichik N, Tarragó T, Burroughs M, Park MS, Itoga NK, Stern HA, Willardson BM and Giralt E. NMR analysis of G-protein betagamma subunit complexes reveals a dynamic G(alpha)-Gbetagamma subunit interface and multiple protein recognition modes. *Proc Natl Acad Sci USA*, Epub Dec 16 (2009)

Takayama K, Tadokoro A, Pujals S, Nakase I, Giralt E and Futaki S. Novel system to achieve one-pot modification of cargo molecules with oligoarginine vectors for intracellular delivery. *Bioconjug Chem*, **20**(2), 249-57 (2009)

Tarragó T, Claasen B, Kichik N, Rodríguez-Mias RA, Gairí M and Giralt E. A cost-effective labeling strategy for the NMR study of large proteins: selective 15N-labeling of the tryptophan side chains of prolyl oligopeptidase. *Chembiochem*, **10**(17), 2736-39 (2009)

Tarragó T, Martín-Benito J, Sabidó E, Claasen B, Madurga S, Gairí M, Valpuesta JM and Giralt E. A new side opening on prolyl oligopeptidase revealed by electron microscopy. *FEBS Lett*, **583**(20), 3344-48 (2009)

Teixidó M, Zurita E, Prades R, Tarragó T and Giralt E. A novel family of diketopiperazines as a tool for the study of transport across the blood-brain barrier (BBB) and their potential use as BBB-shuttles. *Adv Exp Med Biol*, **611**, 227-28 (2009)

Research networks and grants

Collaborations

Activity-based proteomics
Benjamin F Cravatt, The Scripps Research Institute (La Jolla, USA)

Amyloid recycling
Christopher M Dobson, Cambridge University (Cambridge, UK)

Anàlisi conformacional de tres mostres problema d'hormona de creixement (hGH) mitjançant la utilització de tècniques de DC, RMN i LS
Ipsen Pharma SA (Barcelona, Spain)

Applications of the Suzuki reaction to the synthesis of conformationally contrained peptides
Paul-Lloyd Williams, Organic Chemistry Department, University of Barcelona (Barcelona, Spain)

Cell-penetrating peptides
Shiroh Futaki, Kyoto University (Kyoto, Japan)

Design, synthesis and study of P53 ligands
Javier de Mendoza, Institute of Chemical Research of Catalonia (Tarragona, Spain)

Ion-mobility mass spectrometry
Carol Robinson, Oxford University (Oxford, UK)

Nanoparticles for drug delivery
José Luis Pedraz, University of the Basque Country (Vitoria, Spain)

Remote manipulation of protein aggregation
Marcelo Kogan, University of Chile (Santiago, Chile)

Synthesis and conformational analysis of cyclodepsipeptides from marin origin
Fernando Albericio, IRB Barcelona (Barcelona, Spain)

Synthesis and structural studies of β-peptides
Rosa Mª Ortuño, Chemistry Department, Autonomous University of Barcelona (Barcelona, Spain)