View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Nature Precedings

An albumin-derived peptide scaffold capable of binding and catalysis

Immacolata Luisi[1-2], Silvia Pavan[1], Giampaolo Fontanive[1], Alessandro Tossi[2], Fabio Benedetti[1], Adriano Savoini[3], Elisa Maurizio[2], Riccardo Sgarra[2], Daniele Sblattero[4]*, Federico Berti[1]*

[1]Dipartimento di Scienze Chimiche e Farmaceutiche, Università di Trieste, via Giorgieri 1, 34127 Trieste, Italy.

[2]Dipartimento di Scienze della Vita, Università di Trieste, via Weiss 2, 34128 Trieste, Italy.

[3]T&B Associati srl., Area Science Park, Padriciano 99, 34149, Trieste, Italy.

[4]Dipartimento di Scienze della Salute, Università del Piemonte Orientale "Amedeo Avogadro", via Solaroli 17, 28100 Novara, Italy.

Dedicated to the memory of Marta Licciulli, molecular biologist (1976-2012)

Nature Precedings : hdl:10101/npre.2012.6894.1 : Posted 17 Feb 2012

*** Corresponding Authors**

fberti@units.it

daniele.sblattero@med.unipmn.it

Abstract

We have identified a 101-amino-acid polypeptide derived from the sequence surrounding the IIA binding site of human albumin. The polypeptide contains residues that make contact with ligands as warfarin in the parent protein, and eight cysteine residues to form disulfide bridges, which stabilize the polypeptide structure. Seventy-four amino acids are located in six α-helical regions, with the remaining amino acids forming six connecting coil/loop regions. Codon usage optimization was used to express a GST fusion protein in *E. coli* in yields as high as 4 mg/l. This fusion protein retains its structural integrity and aldolase activity, the ability to direct the stereochemical outcome of a diketone reduction, and its binding capacity to warfarin and efavirenz. Notably, this newly cloned polypeptide represents a valuable starting point for the construction of libraries of binders and catalysts with improved proficiency.

Nature Precedings : hdl:10101/npre.2012.6894.1 : Posted 17 Feb 2012

Nature Precedings : hdl:10101/npre.2012.6894.1 : Posted 17 Feb 2012

Introduction

In recent years, a number of different binding proteins have been proposed as alternatives to conventional antibody-based technologies. Affibodies[1], anticalins[2], knottins[3] and neocarzinostatin-based[4] peptide scaffolds are successful examples of such protein-protein or protein-peptide recognition systems[5-7]. However, the recognition of small molecules by peptides and proteins is more demanding. The engineering of novel protein scaffolds for binding small molecules has several drawbacks common also to competing antibody technologies. Both recombinant anti-hapten antibodies selected from in vitro display technologies and artificial protein receptors have, in general, a moderate affinity for small cognate molecules with K_d values in the micromolar range, comparable to those of the primary IgM immune response in animals. Obtaining higher-affinity antibodies or artificial receptors against haptens is still a challenge. Multiple "affinity maturation" approaches have been attempted[8]. As a general strategy, large, full-length gene libraries or, alternatively, focused optimized scaffolds can be constructed. In both cases, specific selection protocols must be devised, using modified haptens to optimize presentation to the immune system or to synthetic libraries. Usually, small haptens are chemically modified at one functional group to introduce a reactive linker suitable for conjugation to a carrier protein or immobilization on a solid surface, often via biotinylation. This modification reduces the number of available interactions and constrains the orientation of the small ligand with respect to potential receptors. Furthermore, the linker and/or the carrier system can elicit a high level of cross reactivity lowering the affinity for the free, unmodified hapten. A detailed example has been reported involving a testosterone-binding peptide derived from neocarzinostatin[9].

Peptide scaffolds, intermediate between large enzymes and small organocatalysts, have also been proposed as tailor-made catalysts for organic reactions[10]. In most cases, however, successful catalysts are synthetic oligopeptides with notably short sequences of amino acids that only play an ancillary role to the main organocatalytic residue side chain (often acting by covalent catalysis) without forming a site for selective substrate recognition[10]. Tanaka and Barbas used phage display to obtain libraries of more structured peptides exhibiting aldolase activity [11,12]. However, a general strategy for developing libraries of structured catalytic peptides that can control regio- and stereoselectivity in a wider variety of reactions has not yet been defined.

Nature Precedings : hdl:10101/npre.2012.6894.1 : Posted 17 Feb 2012

Being interested in the development of artificial peptides that can efficiently combine binding and catalysis, we reasoned that a convenient starting point would be a "wild-type" sequence that is capable of both recognizing a broad spectrum of small molecules with micromolar affinity and catalyzing a wide range of reactions with acceptable proficiency. This polypeptide should ideally present structural features that allow facile detection of binding to the free ligand, avoiding the use of modified targets in the selection step. Such a peptide could be regarded as an "artificial IgM," and the possible development of a library of binders by random or rational mutation strategies could mimic the "in vivo" antibody affinity maturation process in a single step. Serum albumin is a promising candidate due to its ability to bind a variety of small molecules with micromolar affinity[13], and to control the outcome of certain chemical reactions. Broad binding activity and chemical reactivity are associated with the binding site located in the albumin IIA subdomain[14], also known as Sudlow site I. A lysine residue is present in this binding site in both human and bovine serum albumin (K199 and K222 in HSA and BSA, respectively). In addition to being a site for covalent interactions with drugs such as aspirin and benzylpenicillin[15], this basic lysine residue, which is surrounded by a hydrophobic environment, is responsible for the ability of albumin to behave as an enzyme-like catalyst in reactions such as β-eliminations[16], the decomposition of Meisenheimer adducts[17], and the Kemp elimination[18,19]. We have recently reported on albumin's ability to direct the stereoselective reduction of diketones and catalyze the aldol reaction[20-22]. Furthermore, a tryptophan residue (W214) located at the bottom of the main hydrophobic subsite can be exploited as a useful internal fluorescent sensor for binding of this site to different ligands[23,24]. In addition, this residue allows the selection of binders in rapid screening protocols that do not require labeling or immobilization of the target ligand. In this report we describe the selection, cloning and characterization of a 101-amino-acid polypeptide derived from the sequence of the IIA binding site of human albumin that retains the ability of native albumin to bind typical ligands, such as warfarin and efavirenz, and to catalyze or control chemical reactions such as the aldol reaction and the reduction of diketones. These results open the way to the use of this peptide as the base sequence for building libraries of stable artificial receptors and catalysts that can be easily selected for sensing and synthetic applications.

Nature Precedings : hdl:10101/npre.2012.6894.1 : Posted 17 Feb 2012

Results

Identification, cloning and production of the GST-HSA100 polypeptide

Based on structural and functional analyses, we have identified a peptide corresponding to a 101-residue stretch of the human serum albumin sequence (A194 to E294) named HSA100 (Fig. **1a**). This fragment corresponded to approximately half of albumin's subdomain IIa, containing Sudlow site I. This binding site is formed by a continuous sequence of amino acids with no contribution from residues from other subdomains. The HSA100 sequence included all the residues that contact typical albumin ligands, such as warfarin[25], as well as all eight cysteine residues that are involved in the formation of the four disulfide bridges (C200-C246, C254-C252, C265-C279, C278-289), which probably stabilize the polypeptide structure. Seventy-four amino acids were located in six α-helical regions (α_1, Q196-F206; α_2, R209-R222; α_3, A229-H247; α_4, L250-S270; α_5, L284-G292; α_6, L284-G292), and the remaining residues were located in six connecting coil/loop regions. A molecular dynamics simulation was performed at room temperature using the Gromacs package and demonstrated that HSA100 was conformationally stable for over 1000 ns simulation. However, significant unfolding occurred upon the exclusion of C200 or C289 from the designed peptide.

To simplify production of the HSA100 fragment we decided to clone and express its coding sequence in *E. coli*. A preliminary evaluation of protein production was performed by fusing the HSA fragment to either a carrier maltose-binding protein (MBP) or different tag sequences (HIS_6 and strep tags). Even though HSA100 was expressed, purified, and active, the protein yield was extremely poor (data not shown). To improve the yield, we used the following approach: the HSA100 DNA sequence was modified using a codon usage optimization strategy for *E. coli* expression; the fragment was cloned as a fusion protein with a glutathione S-transferase (GST) carrier (Fig. **1b**), generating the GST-HSA100 fusion. After careful optimization of cell culture conditions, we managed to produce 15-20 mg/l of GST-HSA100 fusion protein and obtain up to 4 mg/l of soluble protein. The soluble fraction was purified by affinity chromatography using a GSH resin, resulting in highly pure, stable, full-length protein (Fig. **1c**).

Structural characterization of GST-HSA100

LC-MS/MS and peptide sequencing analyses were performed to confirm the identity of GST-HSA100. Purified GST-HSA100 was trypsin-digested, and the resulting peptides were separated by RP-HPLC. The identities of the peptides were confirmed by comparison of theoretical and experimental *m/z* values of

full-length peptides (MS) and collision-induced dissociation (CID) fragments (MS/MS). The tryptic peptide map referring to the GST-HSA100 sequence is shown in Figure **1d**, and the mass spectrometry data are summarized in supplementary Table 1.

The circular dichroism spectra of the fusion protein in the near- and far-UV regions (Figure **2)** were recorded at neutral pH and temperatures ranging from 25 to 75°C. The far-UV spectrum (Fig. **2a**) was deconvoluted using the convex constraint algorithm[26] to reveal a protein composition of 50% α-helix, 30% coil, 16% turn and <5% β-sheet. Thermal denaturation occurred above 60°C, but at 75°C the conserved amount of α-helix was still over 40%, with no clear trend towards complete unfolding (Fig. **2a**). The presence of disulfide bonds was probed by near-UV CD spectroscopy (Fig. **2b**). Several peaks were identified in the fine structure: the shoulder at 267 nm and minima at 276, 284 and 293 nm resemble a set of peaks attributed to the presence of disulfide bridges in the chiral environment of domain II of HSA[27].

Nature Precedings : hdl:10101/npre.2012.6894.1 : Posted 17 Feb 2012

Small molecule binding activity of GST-HSA100

(±)-Warfarin and efavirenz (Figure **3a**) were chosen as probes to test if the binding capacity of GST-HSA100 was maintained. Warfarin is perhaps the most typical ligand for the subdomain IIa site, and the structure of the HSA complexes formed with both its enantiomers, which have similar affinity, have been fully described[25]. Efavirenz, an inhibitor of HIV reverse transcriptase, is a widely used drug in AIDS therapy and has led to increased interest in the therapeutic drug monitoring of antiretroviral drugs. Efavirenz, as many drugs, binds to HSA and interacts with the subdomain II site[28]. The interactions of the two ligands with the construct were first studied by fluorescence quenching experiments monitoring tryptophan emission. The emission spectrum of apo-GST-HSA100 is shown in Figure **3b** (spectrum 1) as a synchronous scan at Δ = 60 nm and shows an overall maximum at 281 nm excitation (i.e., 341 nm emission), while the emission maximum of native HSA is at 340 nm[27]. Conversely, it is well known that the four GST tryptophan residues have poor emission in native GST with a maximum at 335 nm as they are buried in a hydrophobic environment[29].

Quenching of GST-HSA100 fluorescence (1 μM in phosphate buffer, pH 7.4) was observed in the presence of both warfarin and efavirenz. The effect of efavirenz on the synchronous emission spectrum is shown in Fig. **3b**, and similar results were obtained with warfarin (supplementary Figure 1). Quenching occurred and was accompanied by a blueshift of the excitation maximum from 281 to 278 nm. The highest quenching was obtained at 200 μM efavirenz, which is close to its solubility limit. In the difference

spectrum between unquenched and fully quenched GST-HSA100 emission, there is a shoulder corresponding to an emission maximum at 330 nm. The blueshift from 341 nm upon titration with efavirenz is consistent with a change of the polarity of the environment surrounding the emitter Trp residue, resulting from replacement of the solvent in the active site by the less polar ligand. The buried, inaccessible, tryptophan residues in GST are likely to significantly contribute to the residual unquenched emission observed at 200 μM efavirenz (Fig. **3b**). The titration of fluorescence emission as a function of ligand concentration for HSA, GST-HSA100 and GST is reported in Figure **3c**. Both ligands had a similar quenching effect on HSA and GST-HSA100 , whereas the effect on GST fluorescence was negligible (Fig. **3c**). To determine whether the observed quenching was due to binding or collisional phenomena a Stern-Volmer analysis of the quenching data was performed in the 4300–24500 l mol^{-1} range (Table 1). Assuming a 5-ns decay time for tryptophan fluorescence, the apparent bimolecular quenching constants derived from the Stern-Volmer constants were as high as 8 x 10^{11} – 4.9 x 10^{12} l $mol^{-1}s^{-1}$ (Table 1) and 2-3 orders of magnitude higher than the limit for diffusion–limited collisional quenching. The latter can thus be excluded in favour of static quenching originating from the association of the fluorophore and quenchers in a bimolecular complex[23,30]. The ligand–protein dissociation constants were evaluated by a Hill analysis of the fluorescence data (Table 1). The K_D values obtained for HSA and GST-HSA100 were similar and close to the literature values of 3.7–3.5 μM and 110 μM for warfarin[31,32] and efavirenz binding to native HSA[28], respectively.

The binding of efavirenz to GST-HSA100 was also confirmed by a surface plasmon resonance experiment carried out on a Biacore instrument (Supplementary Fig. 2). Efavirenz was modified at position 1 of the benzoxazin-2-one system to obtain derivative **2b** (Figure **3a**) and this amine was immobilized on a carboxymethylated dextran matrix attached to the gold chip. The affinity measured from the SPR data was comparable to that obtained by fluorimetry (Table 1).

Chemical reactivity control: diketone reduction

1,3-Diols are naturally occurring compounds and important synthetic intermediates. Their preparation in diastereomerically pure form is a frequent target in organic synthesis, and has been achieved through several approaches, including biomimetic and organocatalytic methods. We found that *anti* 1,3-diol **4** can be obtained, with a diastereoisomeric excess of up to 96%, by the stereoselective reduction of diketone **3** (Figure **4a**) with sodium borohydride in aqueous acetonitrile in the presence of stoichiometric amounts of

Nature Precedings : hdl:10101/npre.2012.6894.1 : Posted 17 Feb 2012

bovine or human albumin (Table 2)[11]. The same reaction, without albumin, is not stereoselective, yielding a 1:1 mixture of *syn* and *anti* diols **4**. We carried out the reduction of diketone **3** in the presence of GST-HSA100 and we found that the level of stereoselectivity was comparable to that obtained with native HSA (Table 2). Complete conversion of the substrate to *anti*-diol **4** takes place without loss of selectivity and the *anti* diol can be easily recovered from the aqueous medium by simple extraction after denaturing the peptide with ethanol.

Chemical reactivity control: aldolase activity

The aldol reaction is one of the most important synthetic tools for the stereoselective formation of carbon-carbon bonds. This reaction is extremely useful when proficient catalytic and control systems are employed. Catalytic versions of this reactions have been proposed using metal catalysys, organocatalysis, and bio- and biomimetic catalysis.[ref] We have recently found that HSA and BSA catalyze the aldol addition of acetone to aldehyde **5** (Figure **4b**) and other aromatic aldehydes with an enzyme-like mechanism. The catalyzed process is three orders of magnitude faster than the uncatalyzed reaction, with Michaelis-Menten constants in the millimolar range[21].

GST-HSA100 behaved similarly to its parent protein (Fig. **5**, Table 2) accelerating the reaction by over 1100-fold. The construct exhibited multiple turnover, allowing more than 100 catalytic cycles and full conversion of a 100-fold excess of aldehyde **5** to aldol **6**.

Discussion

HSA100 is the shortest functional peptide that has ever been derived from the sequence of HSA and, to the best of our knowledge, the first albumin fragment successfully expressed in *E. coli*.

In the past decade, several larger fragments have been described. To define the structural elements required for the formation of the warfarin binding site, Rüker and colleagues prepared five fragments, including domains I, II, I-II, I-IIa and Ib-II[27,33,34].

Similar work was completed by the East group on fragments corresponding to domains I-II and II-III [32].

More recently, a recombinant protein corresponding to domain I-II was designed by Fasano[35] to study the allosteric linkage between the warfarin and heme binding sites in the protein. Three recombinant full domains of HSA have also been prepared as potential drug delivery tools[36]. All the described fragments, including native HSA, have been obtained as secreted proteins by expression in *Pichia pastoris*.

Nature Precedings : hdl:10101/npre.2012.6894.1 : Posted 17 Feb 2012

Nature Precedings : hdl:10101/npre.2012.6894.1 : Posted 17 Feb 2012

Using a combined strategy of optimizing DNA sequences and growth conditions, we were able to express and purify soluble, full-length GST-HSA100 protein in *E. coli*. Mass spectrometry in combination with SDS-PAGE and anti-FLAG immunoblotting unambiguously demonstrated the correct production of a full-length recombinant soluble GST-HSA100 fusion construct (Fig. **1c**, Supplementary Table 1).

Our CD spectra support the correct folding of the GST-HSA100 protein. The secondary structure composition inferred from the CD data is as expected for GST-HSA100, containing the five helix and six coil/loop region of the GST-HSA100 domain and two GST domains with a $\beta\alpha\beta\alpha\beta\beta\alpha$ folding topology along the first 84 residues followed by an extended α-helical domain of 132 amino acids at the C-terminus[29]. The thermal stability of the GST-HSA100 protein is also comparable to those of HSA and GST [37,38], and the formation of disulfide bridges is again consistent with a correct folding process. Four disulfide bridges should be present in the GST-HSA100 region of the construct, although none of the four cysteine residues in GST is believed to be involved in disulfide bond formation, even in the predominant homodimer that forms at micromolar concentrations[29,39,40].

Binding activity is conserved in the GST-HSA100 fragment, and the measured affinities fit well with the hosting capacities of the native protein. The conservation of the binding capacity is consistent with the correct organization of GST-HSA100, thereby enabling the correct formation of the binding site. Based on the fluorescence data, the affinity for free efavirenz in solution is similar to that measured on immobilized efavirenz by the SPR sensor. Efavirenz was immobilized to the sensor surface via its heterocyclic nitrogen, leaving the interaction with the protein and the aromatic/cyclopropyl moieties of the molecule exposed. This is consistent with the shape of the IIa binding site, which contains two hydrophobic subsites that host similar aromatic and hydrophobic groups. To design rapid screening methodology for binding by GST-HSA100, we performed the fluorescence-quenching titration in a 384-microwell plate format using a fluorimetric plate reader equipped with 280±15 nm excitation and 350±15 nm emission filters to detect tryptophan fluorescence. Due to the lower sensitivity of this method, microwell screening required a higher concentration of GST-HSA100 to provide a detectable fluorescence signal. However, the results are comparable with fluorimetric titrations in the spectrofluorimeter, enabling a considerably higher throughput and requiring less than a quarter of the volume (Supplementary Figure 3). This creates the possibility of screening libraries of mutated GST-HSA100 for improved binders.

The control of diketone reduction (Fig. **4**, Table 2) was again consistent with the high degree of structural conservation in GST-HSA100. This reaction requires the recognition of the 1,3-diketone in its enol form,

stabilization by ionizable residues H242 and K199 (native protein numbering) in their neutral and protonated forms, respectively, and the correct shape of the binding site around the enolized diketone to control the diastereoselectivity of the products. The initial chemoselectivity occurs in the reduction of the first (aliphatic) carbonyl, which is exposed to the solvent in the albumin-diketone complex, while the aromatic carbonyl is buried inside the albumin binding site. Subsequent diastereoselectivity occurs in the reduction of the second (aromatic) carbonyl[20]. Thus, the role of albumin is to provide a combination of chemo- and stereocontrol that can be maintained in the shortened polypeptide only if the overall shape of the binding site is substantially conserved.

Experimental evidence supports the hypothesis that the HSA-catalyzed aldol reaction (Fig. **4** and **5**, Table 2) occurs through an enamine intermediate via a mechanism similar to that of type 1 aldolases. Lysine 199 reacts with acetone to produce the acetone enamine, which acts as nucleophile toward the aldehyde substrate. The conservation of aldolase activity in the peptide suggests that the peculiar pK of the key lysine residue is not significantly altered in the shortened protein sequence, indicating a conserved local environment. These data demonstrate the possibility of successfully obtaining aldolase peptides with enhanced efficiency and stereoselectivity from mutated libraries of GST-HSA100.

In summary, GST-HSA100 scaffolds can be used in both engineering and display screening technologies and have the potential to overcome the limitations of conventional recombinant antibodies as library templates.

Nature Precedings : hdl:10101/npre.2012.6894.1 : Posted 17 Feb 2012

Methods

Materials

HSA essentially free of fatty acids (A3782), (±)-warfarin (A2250) and all other reagents were purchased from Sigma–Aldrich. SPR reagents, buffers and the research-grade Sensor Chip CM5 were obtained from GE Healthcare Bio-Sciences. Efavirenz was synthesized as previously reported[41]. Phosphate buffer (10 mM Na_2PO_4, 1.76 mM KH_2PO_4, pH 7.4) was used to perform all CD and fluorescence experiments. PBS was used in SPR analyses.

Cloning of the HSA binding site II

The HSA binding site II domain sequence was designed with a codon optimization strategy for *E. coli* expression and ordered from Invitrogen GeneArt. The optimized sequence was subcloned from the original

pMA vector into the prokaryotic expression vector pGEX 4T-1 (GE Healthcare) to yield GST fusion products. All cloning was verified by sequencing.

GST-HSA100 production and purification

Positive colonies harboring the GST construct were grown, and GST-HSA100 was expressed and purified as previously reported [42]. Briefly, cells were grown to an OD of 0.6 at 600 nm and induced overnight with 0.2 mM IPTG at 20°-25°C. Bacteria were harvested by centrifugation, and the cell pellets were lysed with 10 ml of lysis solution for each gram of bacteria, containing lysis buffer (20 mM Tris pH 8.00, 500 mM NaCl, 0.1% Triton X-100), 400 µg/mL lysozyme, 50 µg/mL DNase, and a protease inhibitor cocktail. After centrifugation, the supernatants were collected and purified by affinity chromatography using Glutathione Sepharose 4B (GE Healthcare) according to the manufacturer's instructions. Fractions were subsequently pooled and dialyzed in phosphate buffer or PBS overnight at 4°C. Protein production and purity was verified by SDS-PAGE followed by Coomassie Blue staining and Western blotting using anti-FLAG antibody.

CD spectroscopy

CD spectra were obtained on a Jasco J-710 spectropolarimeter. Spectra were recorded in a 0.2-cm quartz cuvette at temperatures ranging from 5 to 75°C. Data were collected with a data pitch of 0.2 nm, a scanning speed of 50 nm/min and a bandwidth of 1 nm. Each spectrum was the average of 10 scans. A 1.35 µM stock solution of GST-HSA100 was prepared and diluted to 135 nM for CD measurements.

Fluorescence

Warfarin and efavirenz stock solutions were prepared in acetonitrile. Steady-state fluorescence spectra were acquired at 298 K on a CARY Eclipse (Varian) spectrofluorimeter. Synchronous fluorescence spectra were measured in the 240-320 nm excitation range, and the emission was recorded at Δ = 60 nm. The concentration of protein (HSA, GST, or GST-HSA100) was maintained at 1 µM in 350 µl phosphate buffer, whereas the concentration of warfarin or efavirenz was gradually increased (from 1 to 300 µM and 1 to 2 mM for warfarin and efavirenz, respectively). After the addition of each ligand, the fluorescence intensity at the maximum emission wavelength and the drift of such maximums were measured after equilibrium had been reached (15 min).

SPR analysis

The efavirenz derivative **2b** was dissolved at a 1 mM concentration in 1:10 borate buffer (10 mM disodium tetraborate, 1 M NaCl) with 5% of DMSO. Derivative **2b** was then coupled to Sensor chip CM5 using an EDC/NHS coupling kit in 1:10 HBS-EP buffer [0.1 M HEPES, 1.5 M NaCl, 30 mM EDTA, 0.5% v/v p20

Nature Precedings : hdl:10101/npre.2012.6894.1 : Posted 17 Feb 2012

surfactant] Excess active esters were deactivated with ethanolamine. The coupling reactions were performed using a flow rate of 10 µl/min for 7 min, resulting in densities that were approximately 400± 50 RU on average. Analyses of interactions were carried out with a constant flow rate of 10 µL/min in PBS (53 mM $Na_2HPO_4.7H_20$, 12.5 mM KH_2PO_4, 70 mM NaCl, pH 7.4) as running buffer with an optimal GST-HSA100 concentration of 5 µM. After each cycle, both surfaces were regenerated with 30% acetonitrile in 1 mM NaOH. The reference data from a deactivated flow cell were subtracted from the sensorgrams, and the interaction equilibrium constant was calculated using the BIACORE X100 evaluation software.

HPLC

Analyses of the reduction of diketone **3** and the aldol addition of acetone to aldehyde **5** were performed on a HP1100 series instrument equipped with a C18 Phenomenex Luna column (5 µm, 150 x 4.60 mm) using appropriate mixtures of acetonitrile and water as eluent. A flow rate of 0.5 ml/min was used, and the reaction products were monitored at 214 nm. Diketone **3** and its *anti* and *syn* diols **4** were separated using a 68:32 water/acetonitrile mixture. Aldehyde **5** and its aldol product **6** were separated using a 40:60 water/acetonitrile mixture.

Reduction of diketone 3

A solution (0.5 ml) of diketone **3** (43 µM) in acetonitrile was added to 0.5 ml of 43 µM GST-HSA100 in water, and the resulting mixture was maintained at room temperature for 30 min. A solution (0.5 ml) of $NaBH_4$ (75 µM) in water was then added at 20°C, and the resulting mixture was stirred at the same temperature for 2 h. The mixture was acidified with trifluoroacetic acid (150 µl), and 1 ml ethanol was added. Next, the mixture was filtered with a 0.22-µm filter and analyzed by HPLC.

Aldol reaction

GST-HSA100 (200 µM) in 90 µl of phosphate buffer (0.05 M Na_2HPO_4, 0.5 M NaCl, pH 7.5) was incubated at 37°C for 15 min. A solution of aldehyde **5** (2 µl) in acetonitrile was added for final concentrations ranging from 0.5 to 2 mM followed by the addition of 10 µl of acetone. The samples were subsequently sealed and kept at 37°C. Initial velocities of the aldol reaction were obtained by measuring the concentrations of aldehyde and aldol products by HPLC, while the reactions went to 5% completion.

Acknowledgments

We are grateful to Commissariato del Governo nella Regione Friuli – Venezia Giulia (Fondo Trieste Grant 597/08 “Piccoli peptidi per lo sviluppo di biosensori”) for having funded our research project

Nature Precedings : hdl:10101/npre.2012.6894.1 : Posted 17 Feb 2012

and for grants to IL, SP and GF. We are grateful to Regione Piemonte "Piattaforma ImmOnc" for a research fund to DS.

Author contributions

FBer suggested the use of an albumin fragment and identified the GST-HSA100 region. DS, AS, FB and AT co-designed the research. DS and IL designed the expression aystem, IL carried out protein cloning and purification, solution assays and analyzed data. AT analyzed the CD data. EM and RS carried out and analyzed the MS experiments. Binding experiments were carried out by SP and GF, and their results were analyzed by SP, GF, FBer and FB. SP synthesizied the efarirenz derivatives required for SPR experiments, and carried out the SPR analyses. FBer carried out the catalytic activity tests. DS, AS, AT, FBer and FB conceptualized, analyzed and interpreted all studies. FBer and DS co-wrote the manuscript.

Nature Precedings : hdl:10101/npre.2012.6894.1 : Posted 17 Feb 2012

REFERENCES

1. Nygren, P.A. Alternative binding proteins: affibody binding proteins developed from a small three-helix bundle scaffold. *Febs J* 275, 2668-76 (2008).
2. Skerra, A. Alternative binding proteins: anticalins - harnessing the structural plasticity of the lipocalin ligand pocket to engineer novel binding activities. *Febs J* 275, 2677-83 (2008).
3. Kolmar, H. Alternative binding proteins: biological activity and therapeutic potential of cystine-knot miniproteins. *Febs J* 275, 2684-90 (2008).
4. Heyd, B. et al. In vitro evolution of the binding specificity of neocarzinostatin, an enediyne-binding chromoprotein. *Biochemistry* 42, 5674-83 (2003).
5. Binz, H.K. & Pluckthun, A. Engineered proteins as specific binding reagents. *Curr Opin Biotechnol* 16, 459-69 (2005).
6. Skerra, A. Alternative non-antibody scaffolds for molecular recognition. *Curr Opin Biotechnol* 18, 295-304 (2007).
7. Hey, T., Fiedler, E., Rudolph, R. & Fiedler, M. Artificial, non-antibody binding proteins for pharmaceutical and industrial applications. *Trends Biotechnol* 23, 514-22 (2005).
8. Sheedy, C., MacKenzie, C.R. & Hall, J.C. Isolation and affinity maturation of hapten-specific antibodies. *Biotechnol Adv* 25, 333-52 (2007).
9. Drevelle, A. et al. Structures of in vitro evolved binding sites on neocarzinostatin scaffold reveal unanticipated evolutionary pathways. *J Mol Biol* 358, 455-71 (2006).
10. Revell, J.D. & Wennemers, H. Peptidic catalysts developed by combinatorial screening methods. *Curr Opin Chem Biol* 11, 269-78 (2007).
11. Tanaka, F., Fuller, R. & Barbas, C.F., 3rd. Development of small designer aldolase enzymes: catalytic activity, folding, and substrate specificity. *Biochemistry* 44, 7583-92 (2005).
12. Tanaka, F. & Barbas Iii, C.F. Phage display selection of peptides possessing aldolase activity. *Chemical Communications*, 769-770 (2001).
13. Peters, T.J. *All About Albumin: Biochemistry, Genetics, and Medical Applications*, (Academic Press, New York, 1996).
14. Carter, D.C. & Ho, J.X. Structure of serum albumin. *Adv Protein Chem* 45, 153-203 (1994).
15. Yvon, M. & Wal, J.M. Identification of lysine residue 199 of human serum albumin as a binding site for benzylpenicilloyl groups. *FEBS Lett* 239, 237-40 (1988).
16. Klein, G. & Reymond, J.L. An enantioselective fluorimetric assay for alcohol dehydrogenases using albumin-catalyzed beta-elimination of umbelliferone. *Bioorg Med Chem Lett* 8, 1113-6 (1998).
17. Badalassi, F., Wahler, D., Klein, G., Crotti, P. & Reymond, J.L. A Versatile Periodate-Coupled Fluorogenic Assay for Hydrolytic Enzymes *Angew Chem Int Ed Engl* 39, 4067-4070 (2000).

18. Kikuchi, K., Thorn, S.N. & Hilvert, D. Albumin-Catalyzed Proton Transfer. *Journal of the American Chemical Society* 118, 8184-8185 (1996).
19. Hollfelder, F., Kirby, A.J., Tawfik, D.S., Kikuchi, K. & Hilvert, D. Characterization of Proton-Transfer Catalysis by Serum Albumins. *Journal of the American Chemical Society* 122, 1022-1029 (2000).
20. Berti, F. et al. Albumin-directed stereoselective reduction of 1,3-diketones and beta-hydroxyketones to anti diols. *Org Biomol Chem* 9, 1987-99 (2011).
21. Benedetti, F., Berti, F. & Bidoggia, S. Aldolase activity of serum albumins. *Org Biomol Chem* 9, 4417-20 (2011).
22. Benedetti, F., Berti, F., Donati, I. & Fregonese, M. Albumin-controlled stereoselective reduction of 1,3-diketones to anti-diols. *Chem Commun (Camb)*, 828-9 (2002).
23. Zhang, Q., Ni, Y. & Kokot, S. Molecular spectroscopic studies on the interaction between ractopamine and bovine serum albumin. *J Pharm Biomed Anal* 52, 280-8 (2010).
24. Ding, F., Liu, W., Li, N., Zhang, L. & Sun, Y. Complex of nicosulfuron with human serum albumin: A biophysical study. *Journal of Molecular Structure* 975, 256-264 (2010).
25. Ghuman, J. et al. Structural basis of the drug-binding specificity of human serum albumin. *J Mol Biol* 353, 38-52 (2005).
26. Perczel, A., Park, K. & Fasman, G.D. Analysis of the circular dichroism spectrum of proteins using the convex constraint algorithm: A practical guide. *Analytical Biochemistry* 203, 83-93 (1992).
27. Dockal, M., Carter, D.C. & Ruker, F. Conformational transitions of the three recombinant domains of human serum albumin depending on pH. *J Biol Chem* 275, 3042-50 (2000).
28. Bocedi, A. et al. Binding of anti-HIV drugs to human serum albumin. *IUBMB Life* 56, 609-14 (2004).
29. Lim, K. et al. Three-dimensional structure of Schistosoma japonicum glutathione S-transferase fused with a six-amino acid conserved neutralizing epitope of gp41 from HIV. *Protein Sci* 3, 2233-44 (1994).
30. Lackowicz, J.R. *Principles of Fluorescence Spectroscopy 3rd edition*, (Springer, New York, 2006).
31. Rich, R.L., Day, Y.S., Morton, T.A. & Myszka, D.G. High-resolution and high-throughput protocols for measuring drug/human serum albumin interactions using BIACORE. *Anal Biochem* 296, 197-207 (2001).
32. Twine, S.M. et al. Mechanism of binding of warfarin enantiomers to recombinant domains of human albumin. *Arch Biochem Biophys* 414, 83-90 (2003).
33. Dockal, M., Carter, D.C. & Ruker, F. The three recombinant domains of human serum albumin. Structural characterization and ligand binding properties. *J Biol Chem* 274, 29303-10 (1999).
34. Dockal, M., Chang, M., Carter, D.C. & Ruker, F. Five recombinant fragments of human serum albumin-tools for the characterization of the warfarin binding site. *Protein Sci* 9, 1455-65 (2000).
35. Fanali, G., Pariani, G., Ascenzi, P. & Fasano, M. Allosteric and binding properties of Asp1-Glu382 truncated recombinant human serum albumin--an optical and NMR spectroscopic investigation. *Febs J* 276, 2241-50 (2009).
36. Matsushita, S. et al. Functional analysis of recombinant human serum albumin domains for pharmaceutical applications. *Pharm Res* 21, 1924-32 (2004).
37. Rufer, A.C., Thiebach, L., Baer, K., Klein, H.W. & Hennig, M. X-ray structure of glutathione S-transferase from Schistosoma japonicum in a new crystal form reveals flexibility of the substrate-binding site. *Acta Crystallogr Sect F Struct Biol Cryst Commun* 61, 263-5 (2005).
38. Rezaei Tavirani, M., Moghaddamnia, S.H., Ranjbar, B., Amani, M. & Marashi, S.A. Conformational study of human serum albumin in pre-denaturation temperatures by differential scanning calorimetry, circular dichroism and UV spectroscopy. *J Biochem Mol Biol* 39, 530-6 (2006).
39. Kaplan, W. et al. Conformational stability of pGEX-expressed Schistosoma japonicum glutathione S-transferase: a detoxification enzyme and fusion-protein affinity tag. *Protein Sci* 6, 399-406 (1997).
40. Fabrini, R. et al. Monomer-dimer equilibrium in glutathione transferases: a critical re-examination. *Biochemistry* 48, 10473-82 (2009).
41. Pierce, M.E. et al. Practical Asymmetric Synthesis of Efavirenz (DMP 266), an HIV-1 Reverse Transcriptase Inhibitor. *The Journal of Organic Chemistry* 63, 8536-8543 (1998).
42. Di Niro, R. et al. Profiling the autoantibody repertoire by screening phage-displayed human cDNA libraries. *Methods Mol Biol* 570, 353-69 (200**9).**

Nature Precedings : hdl:10101/npre.2012.6894.1 : Posted 17 Feb 2012

Figure Legends

Figure 1. Structure of the human serum albumin HSA-100 fragment
(**a**) The full-length protein and HSA-100 fragment are shown as white and red ribbons, respectively. The structure was obtained from the Protein Data Bank (1BKE). (**b)** Designed construct bearing the polypeptide GST-HSA100. (**c**) SDS-PAGE and anti-FLAG tag Western blot of purified GST-HSA100 protein. (**d**) Tryptic map of GST-HSA100. Tryptic peptides obtained by GST-HSA100 digestion were analyzed by LC-MS/MS. Only peptides covering the GST-HSA100 sequence are shown. Their identities were assigned on the basis of molecular mass (red bars) or peptide sequence (blue bars).

Figure 2. Circular dichroism of GST-HSA100
(a) Far-UV CD spectra. Solid line, 25°C; dashed line, 60°C; dotted line, 75°C. The spectra were recorded at 0.2 μM protein in 10 mM phosphate buffer, pH 7.4. (**b)** Near-UV CD spectrum of GST-HSA100.

Figure 3. GST-HSA100 small ligand binding
(**a**) Reference albumin IIa site ligands. **(b)** Synchronous (Δ=60 nm) fluorescence spectra of GST-HSA100 1 μM in 10 mM phosphate buffer, pH 7.4. 1, no quencher added; 2, 10 μM efavirenz; 3, 100 μM efavirenz; 4, 200 μM efavirenz; and 1-4, difference between spectrum 1 and 4. **(c)** Titration of fluorescence emission of HSA, GST-HSA100 and GST by warfarin and efavirenz. ●, warfarin-HSA; ▲, warfarin-HSA100; ■, warfarin-GST; ○, efavirenz-HSA; Δ, efavirenz-HSA100; □, efavirenz-GST.

Figure 4. Control of reactivity with GST-HSA100
(a) Diastereoselective reduction of 1,3-diketones. **(b)** Albumin-catalyzed aldol addition.

Figure 5.Aldolase activity of GST-HSA100
Michaelis–Menten plot of the initial velocities for the aldol addition of acetone to aldehyde **5** catalyzed by HSA (●) and GST-HSA100 (○).

Nature Precedings : hdl:10101/npre.2012.6894.1 : Posted 17 Feb 2012

Tables

Table 1. Binding parameters for warfarin and efavirenz

	Warfarin		Efavirenz	
	HSA	**HSA100**	**HSA**	**HSA100**
K_{SV}[a] $Lmol^{-1}$	24350	23870	5750	4310
k_q[b] $Lmol^{-1}s^{-1}$	$4.87x10^{12}$	$4.77x10^{12}$	$1.15x10^{12}$	$8.62x10^{11}$
K_D[c] $molL^{-1}$	$2.47x10^{-6}$	$2.70x10^{-6}$	$1.89x10^{-4}$	$2.61x10^{-4}$
n[d]	1.07	1.02	0.98	0.98
K_{ass}[e] $Lmol^{-1}$			$7.1x10^{3}$	$5.9x10^{3}$

[a] Stern-Volmer quenching association constant. [b] Bimolecular quenching kinetic constant, assuming τ_0=5ns for the tryptophan fluorescence decay. [c] Dissociation constant for the protein-ligand complexes from Hill analysis of the quenching data. [d] number of binding sites per molecule of protein from Hill analysis of the quenching data. [e] SPR steady state association equilibrium constant.

Table 1

Table 2: control of reactivity

Protein	Diastereoselective reduction of diketone **3**		Kinetic parameters (at 37°C) for the uncatalyzed, HSA and GST-HSA100-catalyzed aldol addition of acetone to aldehyde **5**				
	%anti	**%syn**	k_{unc}[a] 10^3 min^{-1}	k_{cat}[b] 10^3 min^{-1}	k_{cat}/k_{unc}	K_M[b] 10^{-3} $molL^{-1}$	$k_{cat}/K_M k_{unc}$ $Lmol^{-1}$
BSA	98	2					
HSA	88	12	0.00804	10.0	1244	2.1	$5.9x10^5$
GST-HSA100	90	10	0.00804	9.4	1170	3.0	$3.9x10^5$

[a] observed pseudo first order value. [b] apparent value in 10% aqueous acetone.

Table 2

Nature Precedings : hdl:10101/npre.2012.6894.1 : Posted 17 Feb 2012

Nature Precedings : hdl:10101/npre.2012.6894.1 : Posted 17 Feb 2012

SUPPLEMENTARY INFORMATIONS

Supplementary figure 1. Synchronous (Δ=60 nm) fluorescence spectra of GST-HSA100 1 μM in 10 mM phosphate buffer, pH 7.4. 1, no quencher added; 2, 60 μM warfarin; 3, 100 μM warfarin; 4, 200 μM warfarin.

Supplementary figure 2. SPR sensogram for binding of GST-HSA100 to an efavirenz-coated surface.

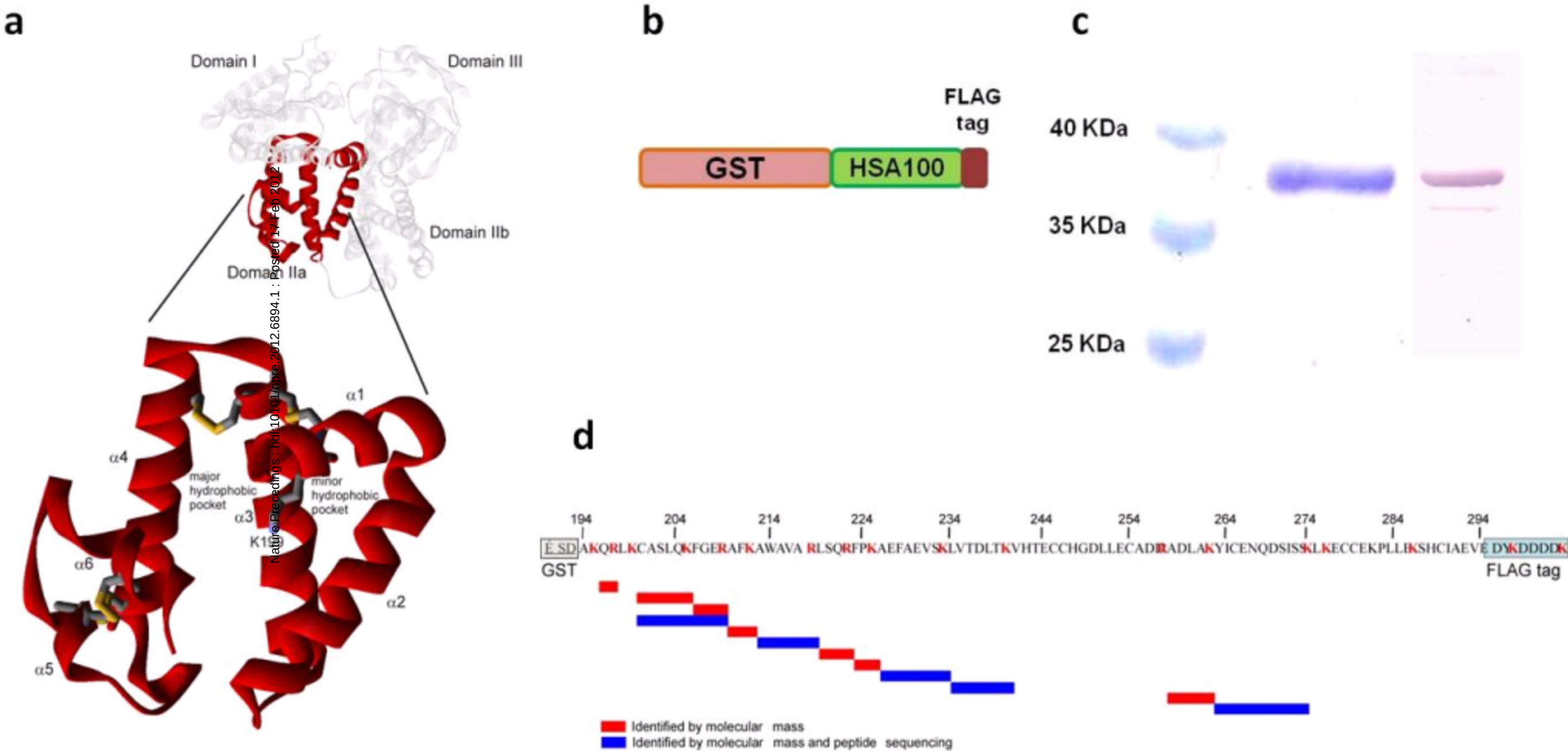
Supplementary figure 3: fluorescence-quenching titration in a 384-microwell plate format using a fluorimetric plate reader equipped with 280±15 nm excitation and 350±15 nm emission filters to detect tryptophan fluorescence. GST-HSA100 was 20 μM.

Table 1: GST-HSA100 peptides detected by mass spectrometry

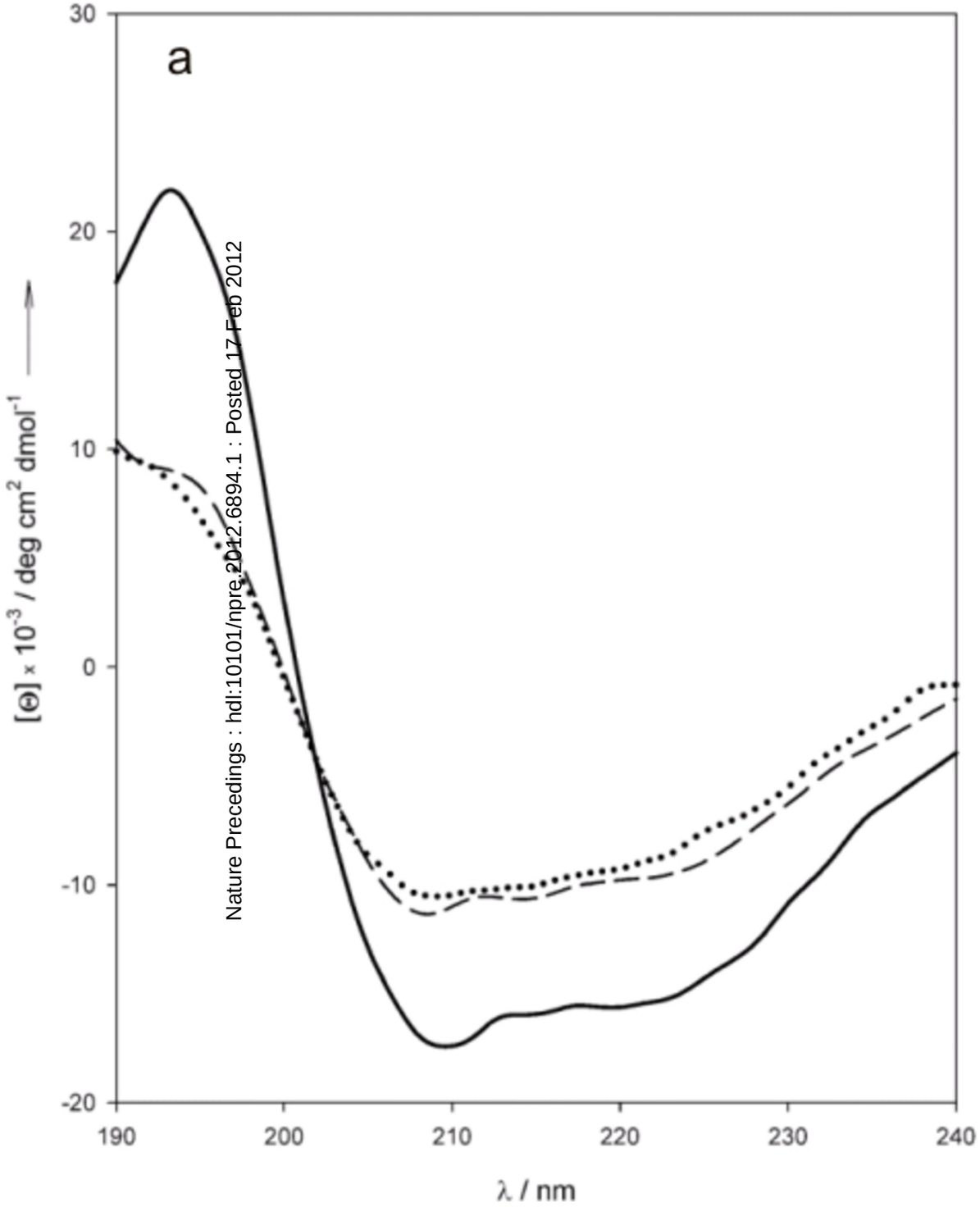
Peptide	Molecular Mass		Elution (min)	m/z^{1+}	m/z^{2+}	m/z^{3+}	sequenced (MS/MS)
	th.	exp.					
196-197	302.17	302.2	3.8	303.2			
200-205	648.32	648.5	11.4	649.3	325.2		
206-209	507.24	507.3	8.5	508.3			
200-209	1137.56	1137.6	19.7	1138.5	569.8	380.2	X
210-212	364.21	364.2	5.1	365.2			
213-218	672.37	672.4	19.9	673.4	337.2		X
219-222	502.28	502.3	4.5	503.3			
223-225	390.22	390.2	10.6	391.2			
226-233	879.43	879.4	19.4	880.4	440.7		X
234-240	788.46	788.5	20.2	789.5	395.3		X
258-262	516.29	516.3	5.9	517.3			
263-274	1385.61	1385.6	20.2		693.8		X

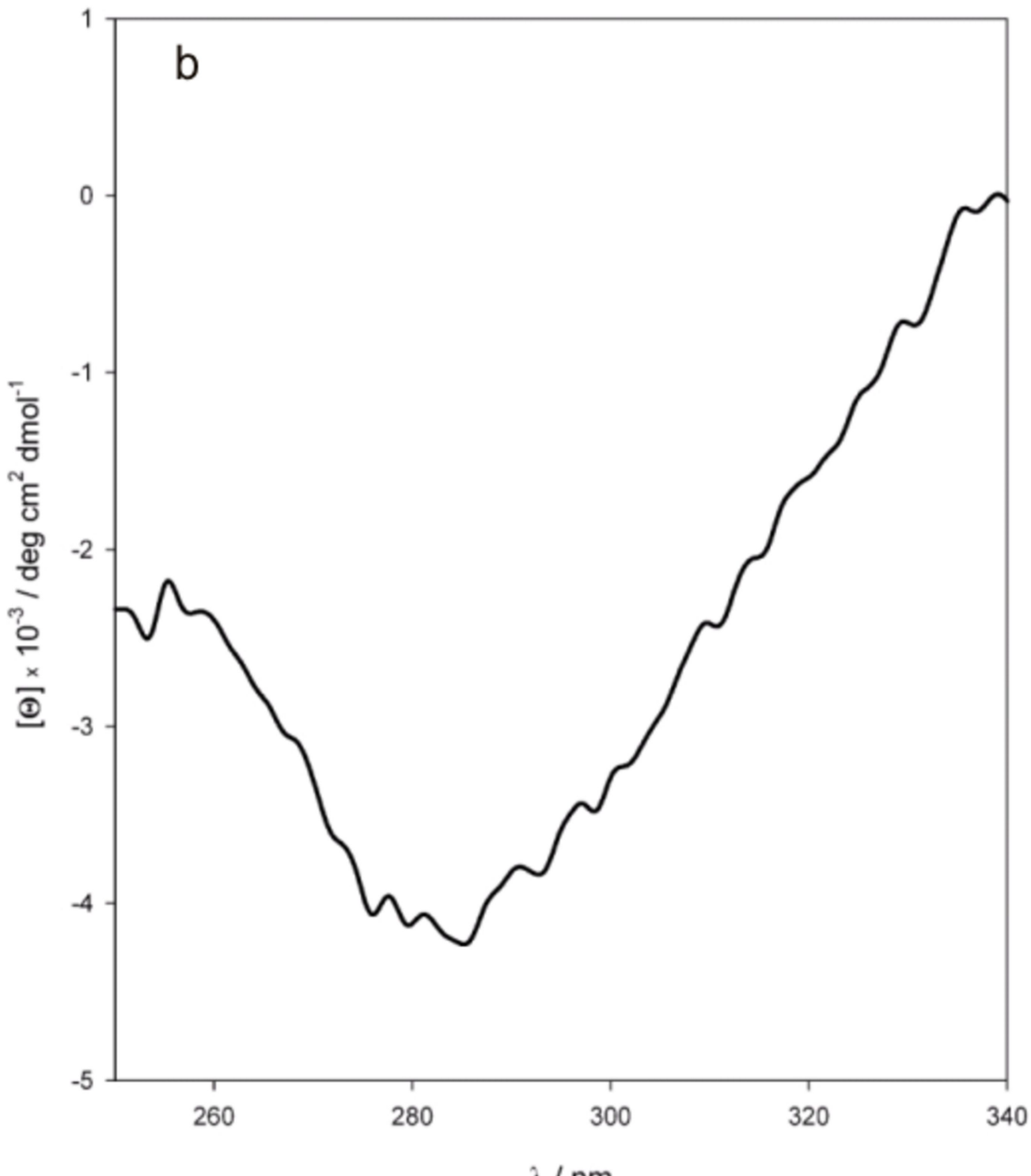
Supplementary table 1 GST-HSA100 LC-MS/MS identification.

Nature Precedings : hdl:10101/npre.2012.6894.1 : Posted 17 Feb 2012



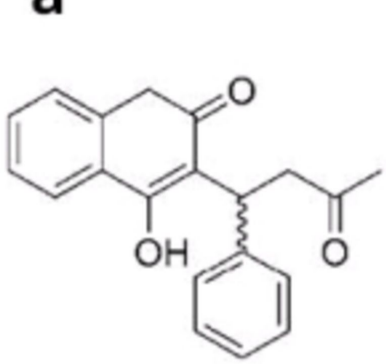
Nature Precedings : hdl:10101/npre.2012.6894.1 : Posted 17 Feb 2012



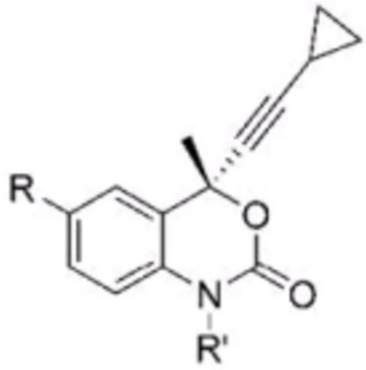


Nature Precedings : hdl:10101/npre.2012.6894.1 : Posted 17 Feb 2012

a



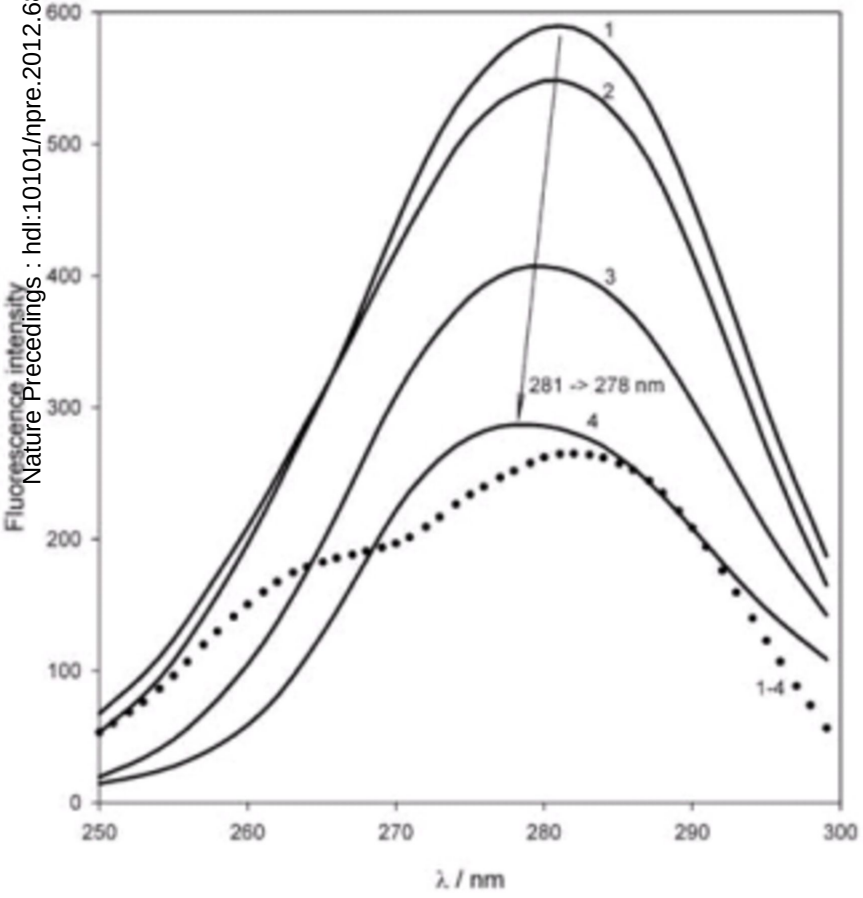
1



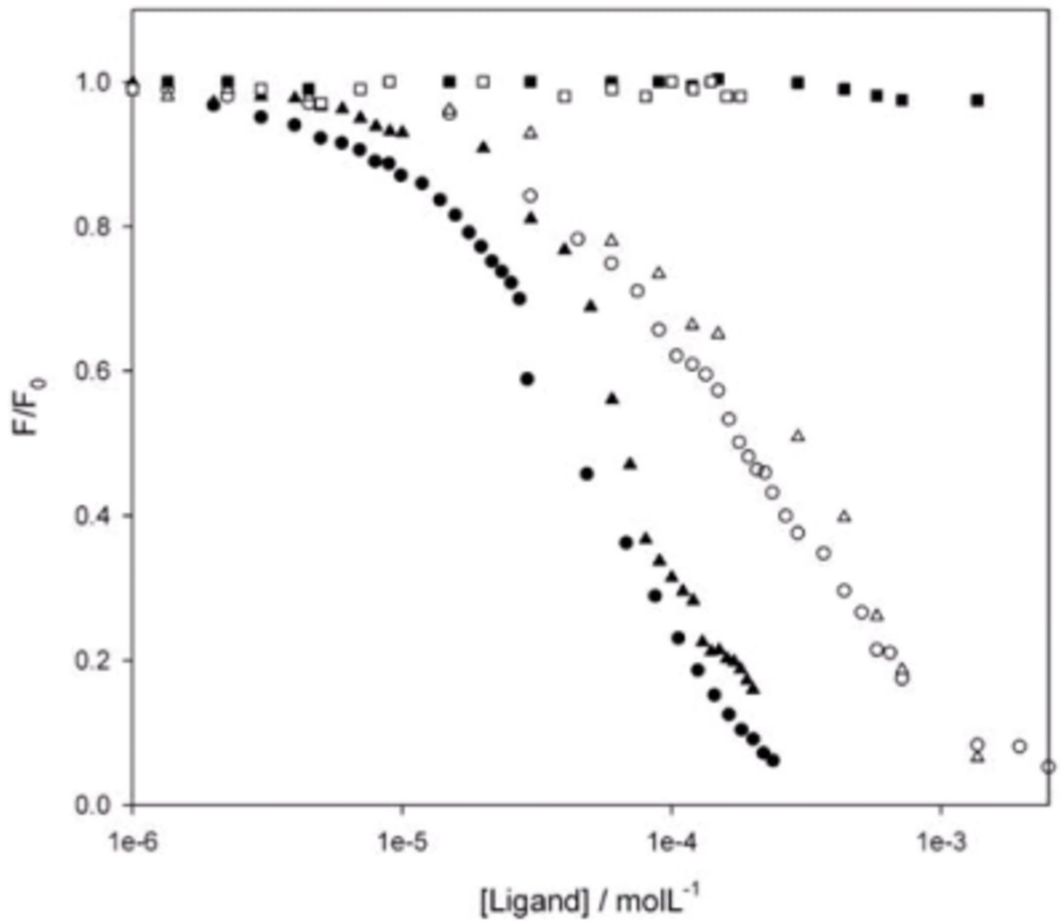
2a-b

a: R=Cl, R'=H
b: R=Cl, R'=$CH_2CONH(CH_2)_6NH_2$

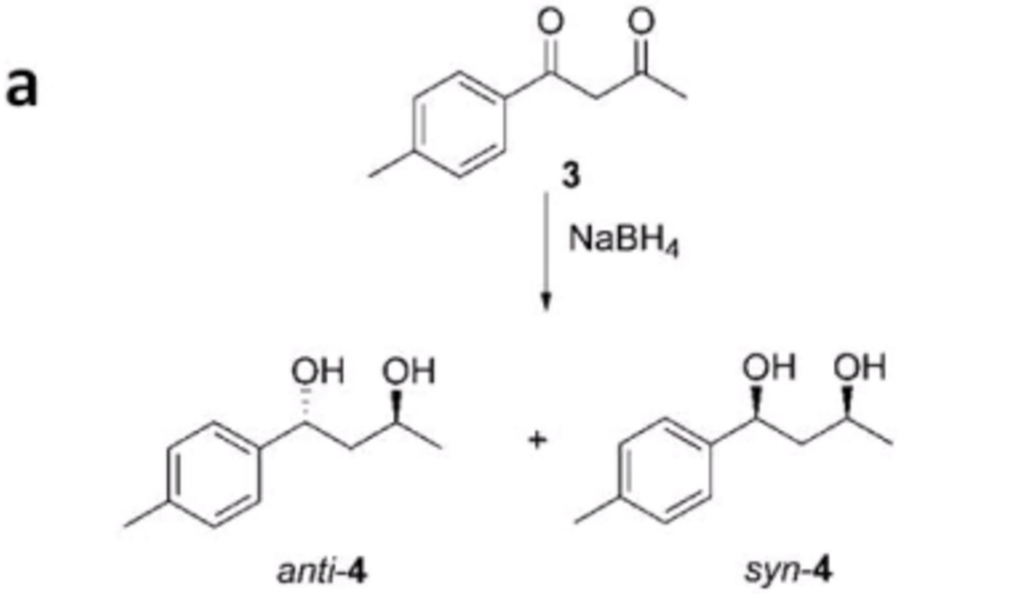
b

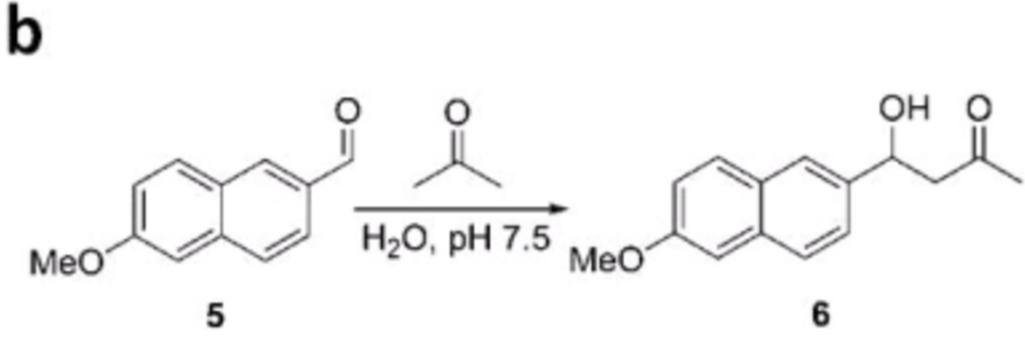


c



Nature Precedings : hdl:10101/npre.2012.6894.1 : Posted 17 Feb 2012





Nature Precedings : hdl:10101/npre.2012.6894.1 : Posted 17 Feb 2012