
NMR STRUCTURAL AND BINDING STUDIES OF RECOMBINANT PROTEINS OF BIOTECHNOLOGICAL INTEREST

Flavia Anna Mercurio

Dottorato in Scienze Biotecnologiche – XXV ciclo
Indirizzo Biotecnologie Industriali e Molecolari
Università di Napoli Federico II



Dottorato in Scienze Biotecnologiche – XXV ciclo
Indirizzo Biotecnologie Industriali e Molecolari
Università di Napoli Federico II



NMR STRUCTURAL AND BINDING STUDIES OF RECOMBINANT PROTEINS OF BIOTECHNOLOGICAL INTEREST

Flavia Anna Mercurio

Dottoranda: Flavia Anna Mercurio

Relatore: Prof. Filomena Rossi

Correlatore: Dr. Marilisa Leone

Coordinatore: Prof. Giovanni Sannia

*Alla mia famiglia,
che mi ha permesso di
raggiungere questo
traguardo.*

INDICE

| | | |
|--|------|----|
| RIASSUNTO | pag. | 1 |
| SUMMARY | pag. | 8 |
| ABBREVIATIONS | pag. | 9 |
| 1. INTRODUCTION | pag. | 11 |
| 1.1. Sam domains | pag. | 11 |
| 1.2. Ephrins and Eph receptors | pag. | 13 |
| 1.2.1. EphA2 receptor | pag. | 15 |
| 1.3. The PI3K effector protein Arap3 | pag. | 17 |
| 1.4. Odin: a member of the Anks family of proteins | pag. | 18 |
| 1.5. Aims of present PhD thesis | pag. | 19 |
| 2. THE NMR SOLUTION STRUCTURE OF THE FIRST SAM DOMAIN OF ODIN AND ITS INTERACTION WITH THE EPHA2 RECEPTOR | pag. | 21 |
| 2.1. BACKGROUND | pag. | 21 |
| 2.2. MATERIALS AND METHODS | pag. | 22 |
| 2.2.1. Protein expression | pag. | 22 |
| 2.2.2. Resonance assignment | pag. | 22 |
| 2.2.3. Relaxation measurements | pag. | 24 |
| 2.2.4. DOSY (Diffusion Ordered Spectroscopy) experiments | pag. | 24 |
| 2.2.5. Structure calculation and analysis of Odin-Sam1 | pag. | 24 |
| 2.2.6. Odin-Sam1/EphA2-Sam binding studies | pag. | 25 |
| 2.2.7. SPR experiments | pag. | 25 |
| 2.2.8. ITC experiments | pag. | 26 |
| 2.2.9. Docking studies | pag. | 26 |
| 2.3. RESULTS AND DISCUSSION | pag. | 27 |
| 2.3.1. Analysis of the aggregation state of Odin-Sam1 | pag. | 27 |
| 2.3.2. NMR solution structure of Odin-Sam1 | pag. | 28 |
| 2.3.3. Odin-Sam1/EphA2-Sam interaction studies | pag. | 29 |

| | | |
|---|------|----|
| 2.3.4. Molecular docking of the Odin-Sam1/EphA2-Sam complex | pag. | 34 |
| 2.4. CONCLUSIONS | pag. | 36 |
| 3. STUDIES OF THE HETEROTYPIC SAM-SAM INTERACTION BETWEEN ODIN AND ARAP3 | pag. | 37 |
| 3.1. BACKGROUND | pag. | 37 |
| 3.2. MATERIALS AND METHODS | pag. | 39 |
| 3.2.1. Protein expression | pag. | 39 |
| 3.2.2. Backbone resonance assignments of the Odin-Sam1/Arap3-Sam complex | pag. | 39 |
| 3.2.3. Relaxation measurements | pag. | 39 |
| 3.2.4. Odin-Sam1/Arap3-Sam binding studies | pag. | 40 |
| 3.2.5. SPR experiments | pag. | 40 |
| 3.2.6. ITC experiments | pag. | 40 |
| 3.2.7. Docking studies | pag. | 41 |
| 3.3. RESULTS AND DISCUSSION | pag. | 42 |
| 3.3.1. Odin-Sam1/Arap3-Sam interaction studies | pag. | 42 |
| 3.3.2. Molecular docking of the Odin-Sam1/Arap3-Sam complex | pag. | 45 |
| 3.4. CONCLUSIONS | pag. | 50 |
| 4. NMR STRUCTURAL CHARACTERIZATION OF THE SAM3 PEPTIDE | pag. | 51 |
| 4.1. BACKGROUND | pag. | 51 |
| 4.2. MATERIALS AND METHODS | pag. | 53 |
| 4.3. RESULTS AND DISCUSSION | pag. | 55 |
| 4.4. CONCLUSIONS | pag. | 59 |
| 5. CONCLUSIONS AND FUTURE PERSPECTIVES | pag. | 61 |
| APPENDIX 1 | pag. | 65 |
| APPENDIX 2 | pag. | 72 |

| | | |
|-------------------------------------|------|----|
| APPENDIX 3 | pag. | 76 |
| ACKNOWLEDGEMENTS | pag. | 79 |
| REFERENCES | pag. | 79 |
| LIST OF PUBLICATIONS | pag. | 91 |
| LIST OF POSTER PRESENTATIONS | pag. | 91 |
| ORAL COMMUNICATIONS | pag. | 91 |
| ATTENDED COURSES AND SCHOOLS | pag. | 92 |
| SCIENTIFIC PRODUCTION | | |

RIASSUNTO

Il mio lavoro di tesi ha riguardato lo studio, mediante tecniche di risonanza magnetica nucleare (NMR) e di modellistica molecolare (*docking*), di alcune proteine di interesse biotecnologico, coinvolte in processi fisio-patologici, contenenti domini Sam (Sterile alpha motif).

Le interazioni proteina-proteina sono essenziali per l'assemblaggio, la regolazione e la localizzazione di complessi proteici nella cellula. L'analisi di tali interazioni a livello molecolare risulta di grande interesse in applicazioni biotecnologiche, in quanto fornisce informazioni utili alla progettazione di sistemi capaci di mimare i siti di interazione, e che quindi possono essere potenzialmente utilizzati in campo terapeutico. Le associazioni tra proteine sono mediate da specifici domini, tra i quali i domini Sam, che sono stati scelti come sistemi modello per gli studi riportati in questa tesi.

I domini Sam sono piccoli moduli proteici di circa 70-80 amminoacidi generalmente strutturati in 5 α -eliche (**Figura 1**), presenti in tutti gli organismi, dal lievito all'uomo (Qiao et al., 2005). Malgrado abbiano strutture molto simili tra loro, i domini Sam presentano un'ampia versatilità relativamente alle interazioni intermolecolari (Kim and Bowie, 2003). Essi possono infatti sia interagire tra loro, formando omo-, eterodimeri, e polimeri (Meruelo and Bowie, 2009), che legare proteine prive di domini Sam (Zhang et al., 2000; Matsuda et al., 2004), l'RNA (Aviv et al., 2003; Green et al., 2003; White et al., 2012) e i lipidi (Barrera et al., 2003; Inoue et al., 2012; Li et al., 2007).

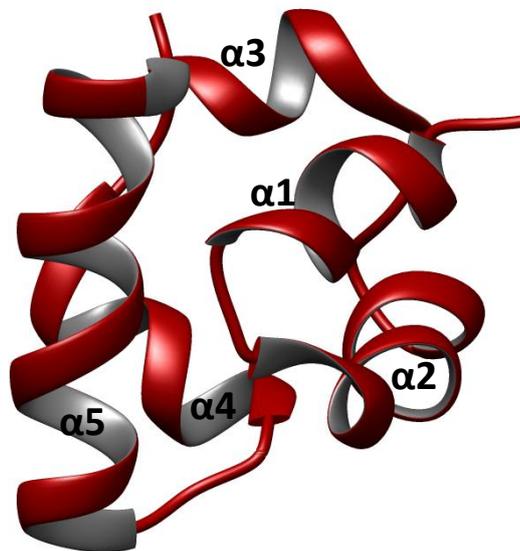


Figura 1. Modello a nastro rappresentante la struttura tridimensionale dei domini Sam.

Il lavoro sperimentale di tesi si è articolato secondo tre fasi successive, descritte di seguito, correlate all'analisi di interazioni tra proteine mediate da domini Sam, e allo studio delle proprietà conformazionali in soluzione di peptidi riproducenti le corrispondenti superfici di legame.

1) Determinazione della struttura in soluzione del primo dominio Sam di Odin e studi di interazione con il dominio Sam del recettore EphA2.

Il recettore tirosina chinasi EphA2 (Ephrin receptor A2) è in grado di regolare importanti processi biologici legati alla motilità cellulare (Egea and Klein, 2007; Himanen et al., 2007; Pasquale, 2005) e svolgere un ruolo fondamentale e controverso in processi patologici quali il cancro (Pasquale, 2008). In particolare, il processo di endocitosi recettoriale, che avviene in seguito al legame con le efrine, e all'attivazione di alcune proteine citosoliche ad attività chinasica che mediano diverse risposte cellulari, è ampiamente studiato per la sua possibile implicazione nei processi tumorali (Pasquale, 2010; Surawska et al., 2004). Infatti, in molti tipi di cancro, in cui EphA2 risulta overespresso, è stata analizzata la potenziale correlazione tra il processo di endocitosi e conseguente degradazione recettoriale, e la diminuzione della malignità tumorale (Pasquale, 2010). La regolazione dell'endocitosi recettoriale in alcuni casi avviene ad opera di proteine che legano EphA2 mediante interazioni eterotipiche di tipo Sam-Sam (Kim et al., 2010; Zhuang et al., 2007), per cui si ipotizza che peptidi o piccole molecole in grado di interferire con la formazione di questi complessi molecolari potrebbero essere potenziali agenti terapeutici.

In particolare, è stato riportato che la fosfatasi lipidica Ship2 (Src homology 2 domain-containing phosphoinositide-5-phosphatase 2) è in grado di inibire il processo di endocitosi di EphA2, con conseguente aumento dell'invasività tumorale e del potenziale metastatico, e che l'interazione tra le due proteine avviene tra i rispettivi domini Sam (Zhuang et al., 2007). In precedenza sono stati riportati studi di interazione tra il dominio Sam di Ship2 (Ship2-Sam) e il dominio Sam di EphA2 (EphA2-Sam), effettuati mediante tecniche NMR e di calorimetria isoterma di titolazione (ITC) (Leone et al., 2008), i quali hanno dimostrato che il complesso EphA2-Sam/Ship2-Sam si forma con un'alta affinità di legame, adottando una topologia comune a diversi altri complessi di tipo Sam-Sam riportati in letteratura (Kurabi et al., 2009; Rajakulendran et al., 2008; Ramachander et al., 2004; Stafford et al., 2011), denominata "Mid-Loop/End-Helix". Questo modello prevede che una delle proteine partecipi all'interazione con la porzione centrale del proprio dominio Sam (Mid-Loop), mentre l'altra opera il legame principalmente mediante l'elica $\alpha 5$ C-terminale ed i *loop* ad essa adiacenti (End-Helix) (**Figura 2**). Inoltre, studi di *docking* molecolare hanno evidenziato che tale complesso è stabilizzato principalmente da interazioni di tipo elettrostatico tra residui carichi positivamente sulla superficie di EphA2-Sam e amminoacidi carichi negativamente della superficie di Ship2-Sam.

Un recente lavoro indica che l'endocitosi recettoriale di EphA2 è regolata anche dalle proteine della famiglia ANKS (Ankyrin repeat and Sam domain containing) (Kim et al., 2010). Questa classe di proteine include Odin e AIDA1b ($A\beta$ PP intracellular domain-associated protein 1B), ed è caratterizzata dalla presenza di due domini Sam in tandem. In particolare il primo dominio Sam di Odin (Odin-Sam1) presenta un'alta omologia di sequenza (circa il 69%) con Ship2-Sam. Da queste premesse è scaturita l'intenzione di analizzare l'eventuale interazione tra Odin-Sam1 ed EphA2-Sam.

Il primo fondamentale passo di questo progetto ha previsto lo studio della struttura in soluzione di Odin-Sam1 mediante tecniche NMR. Al fine di ottenere le assegnazioni delle risonanze per gli atomi H, C e N del *backbone* e delle catene laterali dei singoli amminoacidi della proteina, mediante l'acquisizione di esperimenti NMR bi- e tri-dimensionali, sono stati espressi campioni uniformemente marcati con ^{15}N , e

doppiamente marcati con $^{15}\text{N}/^{13}\text{C}$. In seguito sono stati analizzati esperimenti 3D di tipo NOESY (Nuclear Overhauser Effect Spectroscopy) per ottenere vincoli conformazionali di distanza utili al calcolo della struttura di Odin-Sam1. Tale struttura consiste di 5 α -eliche e rappresenta un canonico ripiegamento di tipo Sam (**Figura 3**).

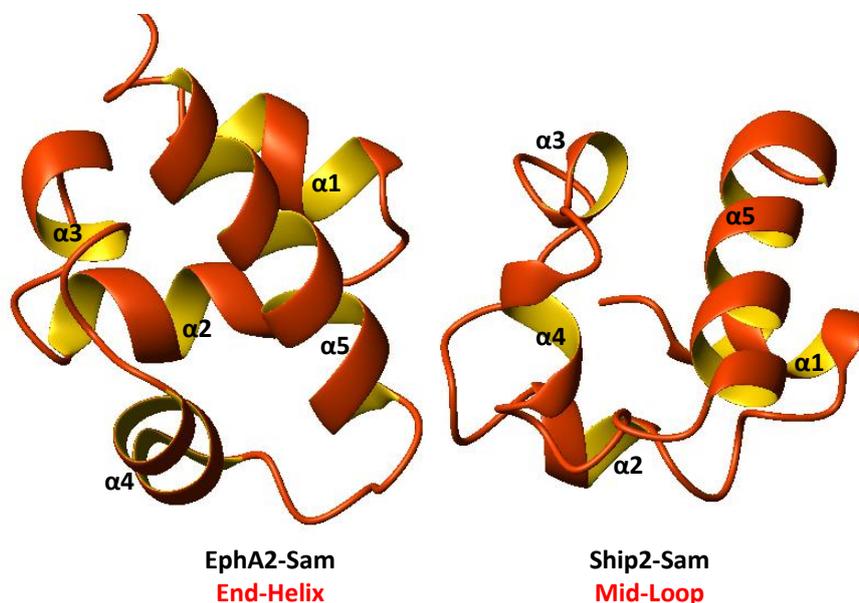


Figura 2. Modello tri-dimensionale del complesso EphA2-Sam/Ship2-Sam (codice PDB: 2KSO (Lee et al., 2011)).

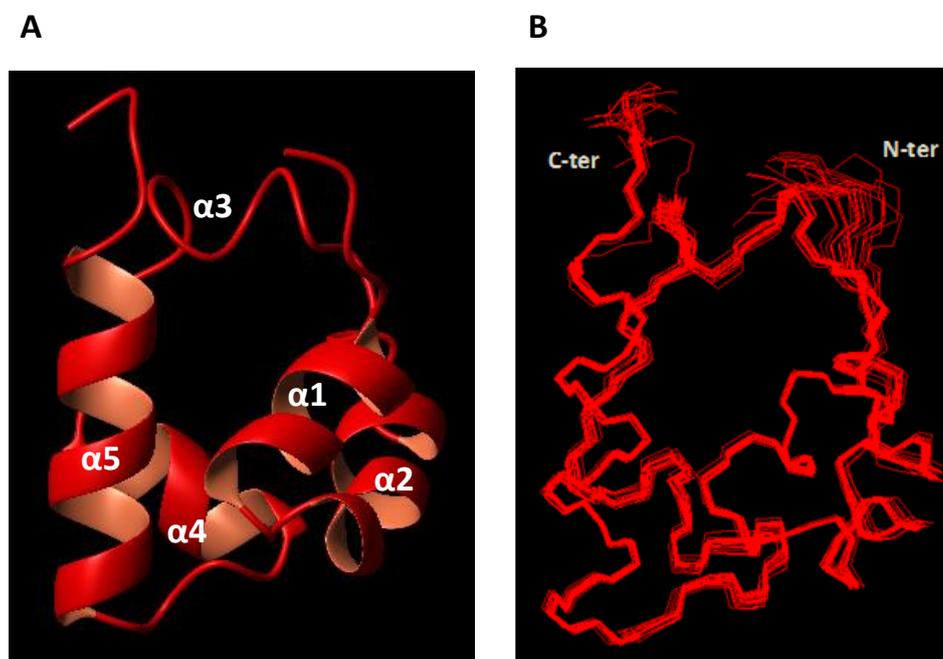


Figura 3. (A) Modello a nastro relativo al primo conformero della struttura NMR di Odin-Sam1 (codice PDB: 2LMR (Mercurio et al., 2012)). **(B)** Rappresentazione delle 20 strutture NMR di Odin-Sam1 sovrapposte sugli atomi del *backbone* dei residui 30-90 (RMSD (Root Mean Square Deviation)= 0.26 ± 0.07 Å).

In seguito, sono stati effettuati studi di interazione con EphA2-Sam. In particolare, tecniche SPR (Risonanza Plasmonica di Superficie) ed ITC hanno fornito misure quantitative riguardanti la formazione del complesso, indicando un' affinità di legame tra le due proteine nell' ordine del basso micromolare. Esperimenti NMR di perturbazione del *chemical shift* (Pellecchia, 2005) sono stati poi condotti e hanno consentito di individuare le reciproche superfici di interazione tra Odin-Sam1 ed EphA2-Sam, e di effettuare ipotesi circa i residui principalmente coinvolti nella formazione del complesso. Questi dati sono stati utilizzati per successivi studi di *docking* molecolare, mediante i quali sono stati costruiti dei modelli tri-dimensionali rappresentanti le possibili modalità di legame tra le proteine. Tali modelli presentano delle similitudini con il complesso EphA2-Sam/Ship2-Sam, e con altri complessi Sam-Sam riportati in letteratura (Kurabi et al., 2009; Rajakulendran et al., 2008; Ramachander et al., 2004; Stafford et al., 2011). Infatti anche in questo caso la topologia di legame è risultata quella "Mid-Loop/End-Helix", in cui Odin-Sam1 fornisce la superficie Mid-Loop che interagisce con l' interfaccia End-Helix di EphA2-Sam, e le interazioni responsabili della formazione del complesso sono principalmente di tipo elettrostatico (**Figura 4**).

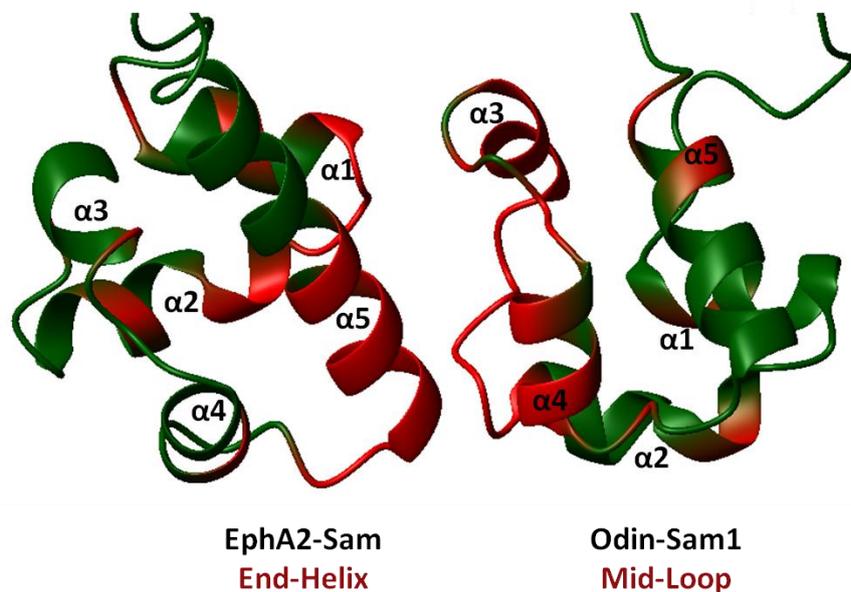


Figura 4. Modello tri-dimensionale rappresentante il complesso EphA2-Sam/Odin-Sam1. Le rispettive superfici di interazione dei domini e le regioni in cui avvengono variazioni conformazionali, determinate mediante studi NMR, sono evidenziate in rosso.

I modelli ottenuti sono stati validati mediante esperimenti di mutagenesi. In particolare è stato progettato un triplo mutante di EphA2-Sam (K38A/R78A/Y81S) (Leone et al., 2008), in cui è stata perturbata la distribuzione di carica nella regione End-Helix mediante opportune sostituzioni amminoacidiche, che presenta solo bassa capacità di legare Odin-Sam1.

Le informazioni strutturali ottenute in questo progetto rappresentano il punto di partenza per una successiva progettazione di molecole di natura peptidica in grado di interferire con la formazione del complesso Odin-Sam1/EphA2-Sam.

2) Studio dell' interazione tra Odin-Sam1 ed Arap3-Sam.

Arap3 (Arf GAP, Rho GAP, Ankyrin repeat and PH domains) è una proteina coinvolta nelle vie di segnalazione della fosfoinositolo-3-chinasi (PI3K) che contiene un dominio Sam nella propria sequenza primaria (Krugmann et al., 2002). È stato riportato che tra le proteine provviste di domini Sam, sia Ship2 che Odin possono interagire con Arap3-Sam (Raaijmakers et al., 2007). La formazione del complesso Ship2-Sam/Arap3-Sam è stata già caratterizzata in dettaglio: l' interazione avviene con alta affinità, mediante topologia "Mid-Loop/End-Helix", ed è mediata da specifici contatti di tipo elettrostatico (Leone et al., 2009). Considerata l' alta omologia di sequenza tra Arap3-Sam ed EphA2-Sam (circa il 58%), abbiamo ritenuto interessante analizzare l' eventuale formazione del complesso Odin-Sam1/Arap3-Sam. A tal fine abbiamo utilizzato metodiche molto simili a quelle adoperate nello studio relativo al complesso Odin-Sam1/EphA2-Sam. In prima analisi, gli esperimenti SPR ed ITC hanno rivelato un' affinità di legame tra le due proteine, nell' ordine del basso micromolare. In seguito, abbiamo ipotizzato quali residui all' interfaccia dei domini siano responsabili dell' interazione mediante esperimenti NMR di perturbazione del *chemical shift* (Pellecchia, 2005). Infine, abbiamo utilizzato i dati ottenuti per costruire dei modelli 3D del complesso Odin-Sam1/Arap3-Sam. Le due proteine interagiscono mediante il modello "Mid-Loop/End-Helix", stabilizzato da interazioni elettrostatiche, comune agli altri complessi qui citati (**Figura 5**).

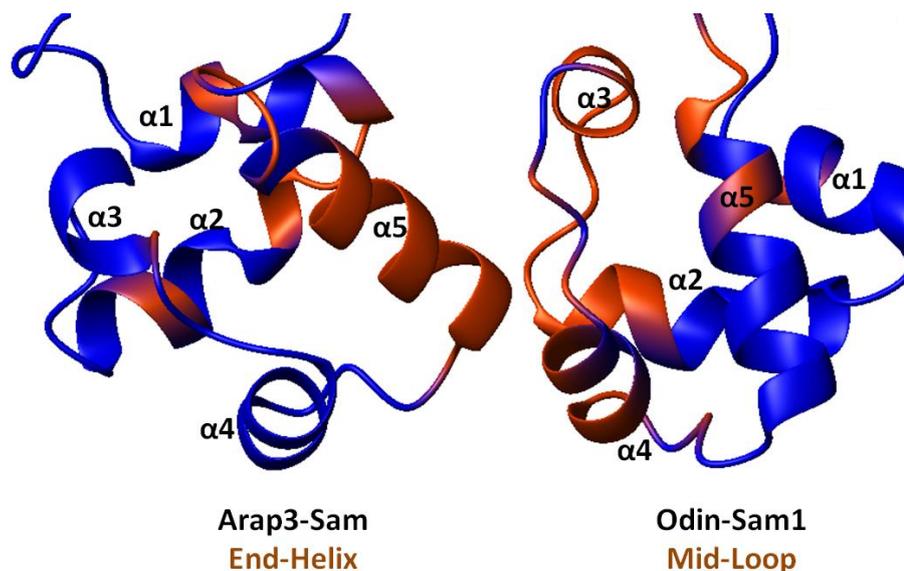


Figura 5. Modello del complesso Arap3-Sam/Odin-Sam1, in cui le interfacce interagenti dei domini e i residui che subiscono variazioni conformazionali in seguito alla formazione del complesso, sono colorate in arancio.

Anche in questa indagine esperimenti di mutagenesi hanno validato le ipotesi formulate riguardo la superficie di interazione di Arap3-Sam nei confronti di Odin-Sam1. Infatti è stato progettato un triplo mutante (H37D, R77D, R80D) (Leone et al., 2009), in cui sono state inserite sostituzioni amminoacidiche nella regione End-Helix di Arap3-Sam, al fine di distruggere interazioni elettrostatiche all' interfaccia del complesso Sam-Sam. Tale proteina mutata perde la capacità di legare Odin-Sam1.

Le informazioni strutturali ottenute in questo secondo studio, saranno utilizzate per generare antagonisti di natura peptidica delle interazioni eterotipiche Sam-Sam coinvolgenti Arap3.

3) Studi strutturali in soluzione del peptide Sam3.

Dopo aver analizzato in dettaglio i fattori strutturali caratteristici dei complessi Odin-Sam1/EphA2-Sam (Paragrafo 1) ed Odin-Sam1/Arap3-Sam (Paragrafo 2), nel gruppo di ricerca in cui ho svolto la mia tesi sono stati progettati diversi peptidi, di dimensioni variabili, includenti le regioni putative di interazione tra domini Sam. Uno di questi peptidi, denominato Sam3, è stato analizzato più nel dettaglio in soluzione mediante tecniche NMR.

Sam3 è un peptide di 43 residui, la cui struttura primaria è riportata in **Figura 6**, riproducente la sequenza della regione 715-758 di Odin-Sam1 (codice UniprotKB Q92625) che include la porzione centrale del dominio Sam di Odin che negli studi precedenti si è ipotizzato fosse implicata nella formazione dei complessi con EphA2-Sam ed Arap3-Sam, insieme all' elica $\alpha 5$ C-terminale (**Figura 6**).

Sam3: Ac-SKLLLNGFDDVHFLGNSVMEEQDLRDIGISDPQHRRKLLQAAR-NH₂

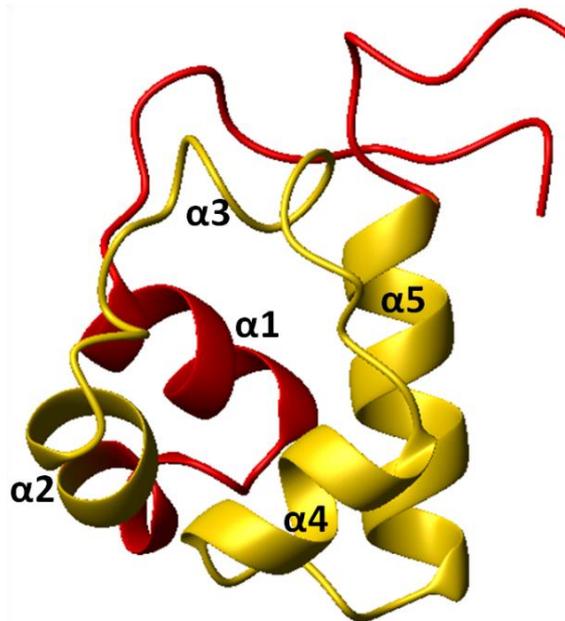


Figura 6. Sequenza primaria di Sam3 e un modello a nastro di Odin-Sam1 con evidenziata in giallo la regione rappresentata dal peptide.

La sequenza corrispondente all' elica $\alpha 5$, esterna alla regione putativa di interazione, è stata inserita per facilitare la strutturazione del peptide in una modalità analoga a quella riscontrata nel dominio Odin-Sam1 intatto.

Sam3 è stato sintetizzato in *batch* utilizzando protocolli *standard* di sintesi in fase solida, purificato con RP-HPLC (Cromatografia Liquida a Fase Inversa ad Alte Prestazioni) e caratterizzato con LC-MS (Cromatografia Liquida-Spettrometria di Massa). Con questo campione sono stati acquisiti spettri NMR bi-dimensionali

protonici utili all' assegnazione delle risonanze dei singoli residui, e all' ottenimento di vincoli conformazionali che saranno utilizzati per il calcolo della struttura tridimensionale. Gli spettri NMR sono stati dapprima acquisiti in tampone fosfato; in tale sistema solvente il peptide risulta piuttosto disordinato e privo di elementi canonici di struttura secondaria. In seguito, ulteriori esperimenti sono stati registrati in presenza di trifluoroetano (TFE), che è di frequente utilizzato quale agente strutturante in studi conformazionali di peptidi (Cammers-Goodwin et al., 1996). I dati raccolti indicano la propensione di Sam3 ad adottare strutture organizzate del tipo elicoidale in presenza di TFE.

Sono in corso studi per valutare la possibile interazione tra Sam3 e i *partner* di legame di Odin-Sam1, ossia EphA2-Sam ed Arap3-Sam.

Conclusioni.

In questa tesi abbiamo intrapreso un approccio interdisciplinare coinvolgente differenti tecniche sperimentali e computazionali (quali: NMR, SPR, ITC, *docking* molecolare, mutagenesi) per studiare in dettaglio la struttura in soluzione e le interazioni intermolecolari di alcuni domini Sam. I risultati ottenuti hanno consentito di ampliare le conoscenze sui dettagli conformazionali responsabili della formazione di complessi eterotipici di tipo Sam-Sam. Sulla base di queste evidenze lo stadio successivo sarà quello di progettare librerie di molecole di natura peptidica e peptidomimetica per applicazioni biotecnologiche, capaci di interferire selettivamente con la formazione dei complessi Odin-Sam1/EphA2-Sam oppure Odin-Sam1/Arap3-Sam. In particolare, riteniamo che antagonisti dell' interazione Odin-Sam1/EphA2-Sam possano modulare l' endocitosi del recettore EphA2, e svolgere un ruolo di agenti terapeutici in malattie quali il cancro. Tali ipotesi saranno validate mediante opportuni saggi in cellula. Per quanto concerne invece il complesso Odin-Sam1/Arap3-Sam, le molecole progettate verranno implementate in saggi cellulari per identificare nuove funzioni biologiche ascrivibili ad Arap3-Sam.

SUMMARY

My PhD thesis has been focused on studies carried out by different experimental and computational techniques, such as NMR (Nuclear Magnetic Resonance), SPR (Surface Plasmon Resonance), ITC (Isothermal Titration Calorimetry), molecular docking, mutagenesis, of Sam (Sterile alpha motif) domains containing proteins that play important roles in physiological or pathological processes.

Protein-protein interactions are essential for the assembly, regulation, and localization of functional protein complexes in the cell. The analysis of these interactions at molecular level is of great interest in many biotechnological fields, as it provides useful information for the design of molecules able to mimic the binding sites and thus, presenting potential therapeutic applications. Protein-protein associations are mediated by specific domains, such as Sam domains.

Within this thesis two heterotypic Sam-Sam interactions were studied in details:

1) the EphA2-Sam/Odin-Sam1 complex; **2)** the Odin-Sam1/Arap3-Sam association.

The tyrosine kinase receptor EphA2 plays a fundamental role in tumorigenesis. The process of EphA2 endocytosis and consequent degradation has been investigated as potential route to reduce tumor malignancy. Odin belongs to the ANKS (Ankyrin repeat domain containing and Sam) protein family; it contains two Sam domains in tandem (Sam1 and Sam2), and is able to regulate EphA2 receptor endocytosis. Instead, Arap3 (Arf GAP, Rho GAP, Ankyrin repeat and PH domains) is a protein involved in the phosphoinositol-3-kinase (PI3K) signaling pathways, which regulates biological processes connected to cell motility.

Firstly, by multidimensional (2D and 3D) NMR methods, the solution structure of Odin-Sam1 was determined. It consists of five α -helices and represents a canonical Sam domain fold. Afterwards, binding studies with EphA2-Sam, and Arap3-Sam were conducted. SPR and ITC experiments revealed a low micromolar binding affinity of Odin-Sam1 for both EphA2-Sam and Arap3-Sam. The reciprocal interaction surfaces of these Sam domains were identified by NMR chemical shift perturbation experiments, and 3D models of the complexes were built by molecular docking techniques. These studies suggested that Odin-Sam1/EphA2-Sam and Odin-Sam1/Arap3-Sam complexes might adopt the canonical Sam-Sam interaction topology called "Mid-Loop/End-Helix", where the central portion of Odin-Sam1 and the C-terminal helix of either Arap3-Sam and EphA2-Sam contribute the binding surfaces.

A peptide, named Sam3, which encompasses the central portion of the Odin-Sam1 involved in complexes formation with EphA2-Sam and Arap3-Sam, together with its C-terminal helix, was synthesized and analyzed by 2D NMR techniques. The peptide reveals unstructured in phosphate buffer, whereas it shows propensity to adopt more ordered helical structures in water/trifluoroethanol mixtures.

Based on the new structural insights gained within my thesis, libraries of molecules (peptides and peptido-mimetics), that could selectively interfere with Odin-Sam1/EphA2-Sam or Odin-Sam1/Arap3-Sam associations, and prove useful in therapeutic and diagnostic applications, will be designed in the near future.

ABBREVIATIONS

Amino Acids: 1 and 3 letters code

| | |
|--------------|---------------|
| A Ala | Alanine |
| N Asn | Asparagine |
| R Arg | Arginine |
| D Asp | Aspartic Acid |
| Q Gln | Glutamine |
| G Gly | Glycine |
| E Glu | Glutamic Acid |
| H His | Histidine |
| I Ile | Isoleucine |
| L Leu | Leucine |
| K Lys | Lysine |
| M Met | Methionine |
| F Phe | Phenylalanine |
| P Pro | Proline |
| S Ser | Serine |
| T Thr | Threonine |
| W Trp | Tryptophan |
| Y Tyr | Tyrosine |
| V Val | Valine |

| | |
|----------------------|---|
| AIDA-1b | A β PP Intracellular Domain-Associated protein 1B |
| ANKS | Ankyrin repeat and Sterile alpha motif domain protein |
| Arap3 | Arf GAP, Rho GAP, Ankyrin repeat and PH domain |
| Arap3-Sam | Sam domain of Arap3 |
| DOSY | Diffusion Ordered Spectroscopy |
| DQFCOSY | Double Quantum-Filtered Correlated Spectroscopy |
| EH | End-Helix |
| EphA2 | Ephrin A2 receptor |
| EphA2-Sam | Sam domain of the EphA2 receptor |
| FPLC | Fast Performance Liquid Chromatography |
| HSQC | Heteronuclear Single Quantum Coherence |
| IPTG | Isopropyl β -D-thiogalactopyranoside |
| ITC | Isothermal Titration Calorimetry |
| K_D | Dissociation Constant |
| KEGG | Kyoto Encyclopedia of Genes and Genomes |
| LB | Luria Broth |
| LC-MS | Liquid Chromatography-Mass Spectrometry |
| ML | Mid-Loop |
| NMR | Nuclear Magnetic Resonance |
| NOE | Nuclear Overhauser Effect |
| NOESY | Nuclear Overhauser Effect Spectroscopy |
| OD | Optical Density |
| Odin-Sam1 | N-terminal Sam domain of Odin |
| Odin-Sam2 | C-terminal Sam domain of Odin |
| PBS | Phosphate Buffer Saline |

| | |
|------------------|---|
| PDB | Protein Data Bank |
| ppb | parts per billion |
| ppm | parts per million |
| PSGE | Pulsed Gradient Spin-Echo |
| RMSD | Root Mean Square Deviation |
| RP-HPLC | Reversed Phase-High Performance Liquid Chromatography |
| RU | Resonance Units |
| Sam | Sterile alfa motif |
| Ship2 | Src homology 2 domain-containing Phosphoinositide-5-phosphatase 2 |
| Ship2-Sam | Sam domain of Ship2 |
| SMART | Simple Modular Architecture Research Tool |
| SPR | Surface Plasmon Resonance |
| TFE | Trifluoroethanol |
| TOCSY | Total Correlation Spectroscopy |
| TSP | Trimethylsilyl Propionic acid |

1. INTRODUCTION

1.1. Sam domains

Sam (Sterile alpha motif) domains were initially identified on the basis of the conservation of a small module made up of approximately 70 amino acids in 14 eukaryotic proteins (Ponting, 1995). The domain was named the Sterile alpha motif domain because of its presence in yeast proteins crucial for sexual differentiation, and the predicted high α helix content (Ponting, 1995; Schultz et al., 1997) (**Figure 1**).

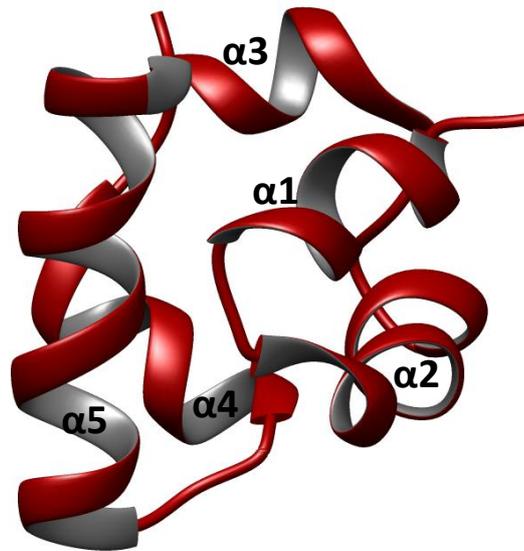


Figure 1. Ribbon representation of the Sam domains 3D structure.

The SMART (Simple Modular Architecture Research Tool) (<http://smart.embl.de/>) (Shultz et al., 1998) database now identifies more than 4000 Sam domain containing proteins in all genomes. These domains are categorized as protein binding modules; however, unlike others protein modules, such as SH2 (Src Homology 2), SH3 (Src Homology 3) and PDZ (Post synaptic density protein, Drosophila disc large tumor suppressor, and Zonula occludens-1 protein) domains, which have conserved functions, Sam domains are not easily categorized (Kim and Bowie, 2003; Qiao and Bowie, 2005).

In fact, although it has been suggested a role in developmental regulation due to the presence of Sam domains in many proteins involved in these processes (Schultz et al., 1996), they just present similar structures but a variety of functions (**Table 1**) (Kim and Bowie, 2003). This behavior is explained by the fact that Sam domains are present in all sub-cellular locations, exist in a wide number of stoichiometries, and are able to bind different partners (Kim and Bowie, 2003).

In particular, Sam domains have been shown to homo-, hetero-oligomerize, and polymerize (Meruelo and Bowie, 2009), forming multiple self-association architectures, and also bind to proteins lacking Sam domains (Matsuda et al., 2004; Zhang et al., 2000), RNA (Aviv et al., 2003; Green et al., 2003; White et al., 2012), and lipids (Barrera et al., 2003; Inoue et al., 2012; Li et al., 2007).

| % proteins involved | Description |
|---------------------|---|
| 43.62 | Axon guidance |
| 7.98 | Phosphatidylinositol signaling system |
| 4.79 | Glycerophospholipid metabolism |
| 4.79 | Glycerolipid metabolism |
| 4.26 | Fc epsilon RI signaling pathway |
| 4.26 | Inositol phosphate metabolism |
| 4.26 | MAPK signaling pathway |
| 4.26 | T cell receptor signaling pathway |
| 4.26 | Natural killer cell mediated cytotoxicity |
| 3.72 | Tight junction |
| 3.72 | p53 signaling pathway |
| 2.66 | Tryptophan metabolism |
| 1.06 | Benzoate degradation via CoA ligation |
| 0.53 | GnRH signaling pathway |
| 0.53 | Apoptosis |
| 0.53 | Olfactory transduction |
| 0.53 | Insulin signaling pathway |
| 0.53 | Taste transduction |
| 0.53 | Wnt signaling pathway |
| 0.53 | Gap junction |
| 0.53 | Calcium signaling pathway |
| 0.53 | Long-term potentiation |
| 0.53 | Melanogenesis |
| 0.53 | Hedgehog signaling pathway |
| 0.53 | Progesterone-mediated oocyte maturation |

Table 1. Information based on mapping of SMART genomic protein database to KEGG (Kyoto Encyclopedia of Genes and Genomes) orthologous groups (<http://www.genome.jp/kegg>) (Kanehisa et al., 2012). KEGG is a database of gene catalogs from completely sequenced genomes which are linked to cell functions. Percentage points are related to the number of proteins with SAM domains which could be assigned to a KEGG orthologous group, and not all proteins containing SAM domain. Proteins can be included in multiple pathways, i.e. the numbers above will not always add up to 100%.

Several Sam domains have been observed to form dimers with other Sam domains in both homo and heterotypic modes, by means of head-to-head (Walker et al., unpublished results), tail-to-tail (Rajakulendran et al., 2008; Stapleton et al., 1999; Thanos et al., 1999), and head-to-tail (Kurabi et al., 2009; Leone et al., 2008; Leone et al., 2009; Mercurio et al., 2012; Mercurio et al., 2013; Rajakulendran et al., 2008; Stafford et al., 2001) orientations (**Figure 2**). For example, the Ephrin B4 kinase (PDB code: 2QKQ- Walker et al., unpublished results) Sam domain has been crystallized in a head-to-head configuration, while the Ephrin B2 (PDB code: 1B4F- Stapleton et al., 1999) receptor forms a homodimer in a tail-to-tail mode. *Drosophila* proteins Ph and Scm bind to each other in a head-to-tail topology (PDB code: 1PK1- Kim et al., 2005), also called “Mid-Loop/End-Helix” (Leone et al., 2008; Leone et al., 2009; Mercurio et al., 2012; Mercurio et al., 2013; Rajakulendran et al., 2008), which is the most conventional interaction mode encountered in Sam-Sam complexes.

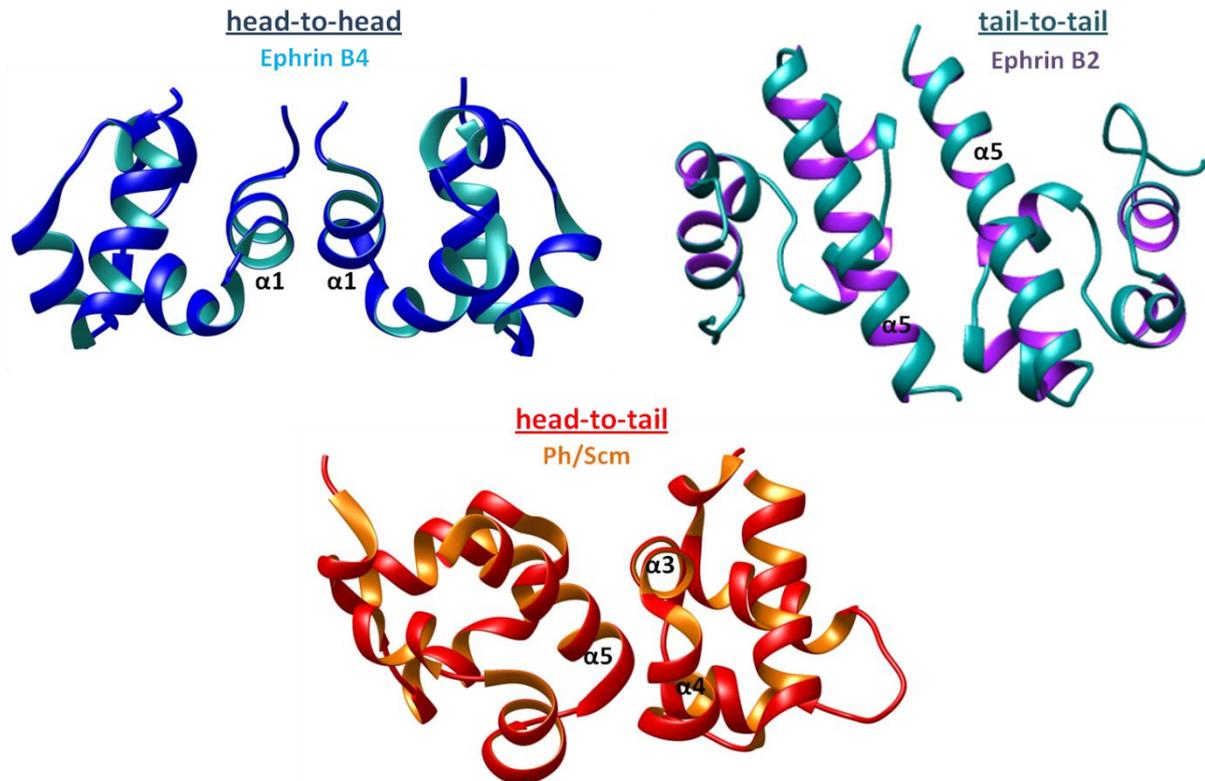


Figure 2. Three examples of different Sam-Sam interactions.

Sam domains often self-associate by forming polymeric structures. For instance, polymeric Sam domains have been characterized in transcriptional repressors (Kim et al., 2001; Kim et al., 2002; Kim et al., 2005; Qiao et al., 2004), scaffolding proteins (Baron et al., 2006; De Rycker et al., 2003), and regulatory enzymes (Harada et al., 2008). Among transcriptional regulators, TEL (Translocation Ets Leukemia) has been frequently linked to human leukemias, and it is thought that these diseases arise because Sam mediate polymerization constitutively activates mitogenic proteins (Carroll et al., 1996; Golub et al., 1996).

Sam domains also appear to possess the ability to bind RNA (Aviv et al. 2003; Green et al., 2003; White et al., 2012). An example is the Sam domain of human SAMHD1 protein, a retroviral HIV restriction factor expressed in myeloid cells, which is dispensable for retroviral restriction, oligomerization, and RNA binding of the protein (White et al., 2012).

In a few cases Sam domains can also bind lipids (Barrera et al., 2003; Inoue et al., 2012; Li et al., 2007). Recently it has been reported that Sam and DDHD domains of A1 KIAA0725p, a member of the intracellular phospholipase A1 family of proteins implicated in organelle biogenesis and membrane trafficking, are essential for protein phosphoinositide binding and phospholipase activity (Inoue et al., 2012).

1.2. Ephrins and Eph receptors

Ephs are the largest family of receptor tyrosine kinases (RTKs) in the animal kingdom (Pitulescu and Adams, 2010). Together with their ligands (ephrins), represent an important cell communication system with well-known roles in

physiological and pathological processes (Pasquale, 2008). In humans Eph receptors are fourteen and, on the basis of sequence homology, have been classified into two groups: EphA (EphA1-8 and EphA10) and EphB (EphB1-4 and EphB6) (Eph-Nomenclature-Committee, 1997). The ephrins have the singularity of being membrane-bound; in particular ephrins belonging to subclass A (ephrin A1-5) are linked to membrane by means of a GPI (Glycosylphosphatidylinositol) linkage, while ephrins of subclass B (ephrin B1-3) are transmembrane proteins (**Figure 3**). EphA receptors typically bind ephrins A, and EphB receptors bind ephrins B (Gale et al, 1996); also limited cross-binding between members of the two classes are reported (Himanen et al., 2004).

The Eph receptors have a ligand binding domain at the N-terminus (Labrador et al., 1997), followed by a central cysteine-rich region, and two fibronectin type III repeats near the membrane-spanning segment (Pasquale, 1997). The cytoplasmatic region of the Eph receptors contains a short juxtamembrane region with several conserved tyrosine residues, followed by a Sam domain, and a C terminal PDZ-binding motif (Pasquale, 1997) (**Figure 3**).

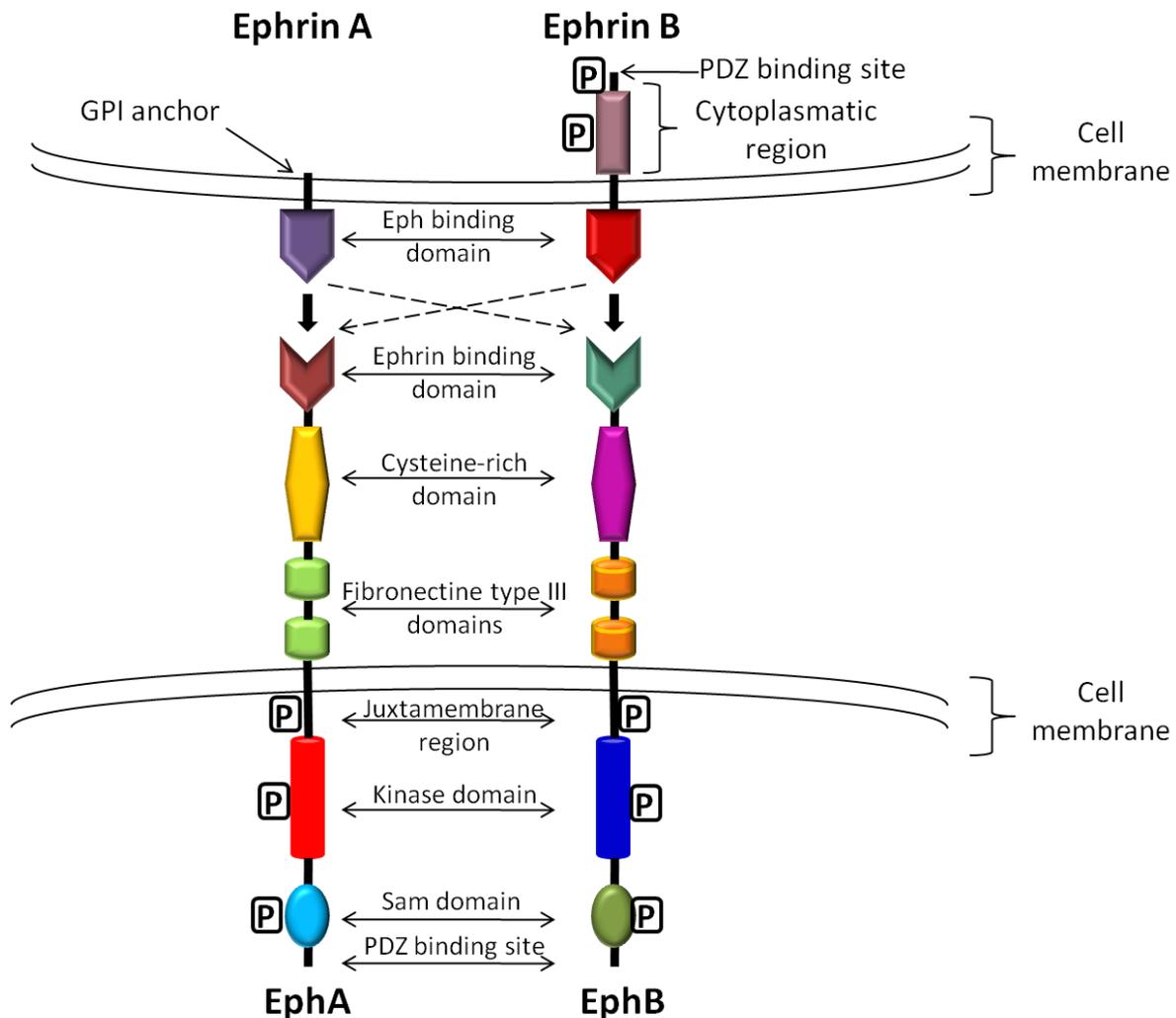


Figure 3. Domain structures of Eph receptors and ephrins, with tyrosine phosphorylation sites indicated by the symbol P.

Generally Eph receptors interact with the cell surface-associated ephrins at sites of cell-cell contacts (Pasquale, 2010). The activation of Eph receptors by means of ephrins gives rise to a bi-directional signal: a forward signal, spreading into receptor-expressing cell, which determines Eph receptor autophosphorylation on specific tyrosine residues, activation of its kinase activity, and association with various effector proteins; and a reverse signal, that propagates in ephrin-expressing cells, which partially depends on tyrosine phosphorylation of the ephrin cytoplasmatic region, and on association with other proteins (Pasquale, 2008). In many cases Eph/ephrin interactions generate repulsive signals that are involved in cell migration and movements in a wide range of processes, such as axon guidance in the nervous system (Wilkinson, 2001), cell positioning in the intestinal epithelium (Battle et al., 2002), and angiogenesis (Adams and Klein, 2000).

A mechanism that terminates Eph/ephrin interactions, with consequent cell separation, is endocytosis (Pitulescu and Adams, 2010). During this process, the intact receptor-ligand complex, together with the surrounding plasma membrane, can be internalized into the Eph- or ephrin- expressing cell (Marston et al., 2003; Zimmer et al., 2003). Some studies suggest that c-Cbl, a RING finger E3 ubiquitin ligase, participates in induction of Eph receptors degradation (Fasen et al., 2008).

Eph receptors direct cell movements by regulating the actin cytoskeleton and the associated cell-matrix adhesions (Miao et al., 2000; Carter et al., 2002). Eph/ephrin signaling is often coupled to regulation of the activity of small Rho family GTPases, such as Rac, Rho, or Cdc42 that connect bidirectional receptor-ligand interactions to changes in the actin cytoskeleton (Groeger and Nobes, 2007; Noren and Pasquale, 2004).

Eph receptors are involved in the pathogenesis of many diseases, above all cancer (Pasquale, 2008). In fact, besides their expression in normal tissues, Eph receptors result present, and often over-expressed, in almost all types of cancer cells (Ireton and Chen, 2005; Noren and Pasquale, 2007), with consequent increasing of angiogenesis and tumor vasculature (Surawska et al., 2004). The up-regulation of Eph receptor is mainly observed in the more aggressive stages of tumor progression (Dodelet and Pasquale, 2000). However, decreased Eph receptor levels have also been reported in a few malignant cancer cell lines (Battle et al., 2005; Hafner et al., 2006; Li et al., 2009). Probably, an initial Eph receptor over-expression, due to oncogenic signaling pathways activation, is followed by an epigenetic silencing in more advanced stages of tumoral progression (Battle et al., 2005). Receptors expression levels can also be regulated by alterations and changes in mRNA stability (Huusko et al., 2004; Winter et al., 2008); indeed, several Eph receptor genes are located in chromosomal regions that are frequently lost in cancer cells (Sulman et al., 1997; Huusko et al., 2004), whereas others Eph genes are located in amplified regions (Kang et al., 2009). Due to the evident connection of Eph receptors with tumoral pathogenesis, they are considered promising therapeutic agents in cancer (Pasquale, 2010).

1.2.1. The EphA2 receptor

The EphA2 receptor (EphA2) is involved in a large number of physiological and pathological processes, which it regulates by means of complex and still poorly understood mechanisms (Surawska et al., 2004).

A few receptor physiological activities are well known: EphA2 stimulation induces activation of RhoA GTPase through a PI3K (Phosphoinositol-3-kinase) dependent pathway, that regulates cell migration (Fang et al., 2005; Miao et al., 2003), and reduces FAK (Focal Adhesion Kinase) phosphorylation, resulting in the inhibition of integrin-mediated cell adhesion (Miao et al., 2000).

Due to the evident EphA2 connection to many types of cancer (such as breast, prostate, lung, ovarian, cervical, esophageal and colorectal cancers, and malignant melanoma (Ireton and Chen, 2005)), this protein is receiving great attention as possible anti-tumoral target (Pasquale, 2010). It has been reported that EphA2 over-expression is sufficient to generate tumorigenesis in mammalian epithelial cells, when EphA2 fails to bind its ligands (ephrins A1-5), whereas receptor stimulation by soluble ephrin A1 reverses growth and invasiveness of EphA2 transformed cells (Zelinski et al., 2001). EphA2 over-expression in cancer does not affect cells proliferation but changes their malignant behaviour, since the receptor is not able to bind its ligands, with a consequent increasing of extracellular matrix adhesion and higher metastatic potential (Ruoslahti, 1999).

EphA2 has also been implicated in tumor invasion of blood vessel, (Brantley et al., 2002). In fact, *in vivo* studies support an important role for EphA2 as regulator of VEGF (Vascular Endothelial Growth Factor)-mediated angiogenesis (Cheng et al., 2002).

A specific process related to EphA receptors, the endocytosis and subsequent degradation, is considered to be critical in regulation of cell repulsive events (Cowan et al., 2005). In the specific case of EphA2, the connection between endocytosis and pro-cancer activities is under investigation; indeed, enhanced receptor internalization and degradation are associated with decreased malignant cell behavior (Zhuang et al., 2007). A few strategies to reduce EphA2 pro-cancer activity by inducing its endocytosis have already been developed, and rely on monoclonal antibodies (Carles-Kinch et al., 2002), ephrin-mimetic peptides (Koolpe et al., 2002), or adenovirus-expressing the ephrin-A1 ligand (Noblitt et al., 2004).

Proteins which normally interact with EphA2 as downstream effectors in cancer cells can affect receptor endocytosis (Kim et al., 2009; Zhuang et al., 2007).

An example is the lipid phosphatase Ship2 (Src homology 2 domain-containing phosphoinositide-5-phosphatase 2), which, together with the PI3K kinase, is supposed to act as a regulator of Rho family GTPase in response to EphA2 activation (Zhuang et al., 2006). Ship2 is provided like EphA2 with a Sam domain. Upon ephrin binding to EphA2, Ship2 is recruited at the receptor site by an heterotypic Sam-Sam association. *In vitro* experiments, performed in MDA-MB-231 mammary carcinoma cells, have demonstrated that Ship2 over-expression increases EphA2 stability by inhibiting its endocytosis and degradation, with a consequent rising of tumor malignancy, while suppression of Ship2 expression by small interfering RNA-mediated gene silencing (siRNA) enhances receptor internalization and degradation (Zhuang et al., 2007).

The NMR (Nuclear Magnetic Resonance) solution structure of the Ship2 Sam domain (Ship2-Sam) (PDB code: 2K4P) was previously reported, together with binding studies with the EphA2 receptor Sam domain (EphA2-Sam) (Leone et al., 2008; Lee et al., 2012). ITC (Isothermal Titration Calorimetry) studies have shown that EphA2-Sam binds Ship2-Sam with a low micromolar binding affinity. Binding studies have indicated that these proteins bind by adopting a topology common to several Sam-

Sam complexes, called "Mid-Loop/End-Helix" (Leone et al., 2008; Lee et al., 2012) (See Paragraph 1.1).

EphA2 endocytosis is also regulated by the ANKS (Ankyrin repeat domain containing and Sam) protein family (Kim et al., 2009). In particular it has been revealed that Odin (See Paragraph 1.4), a member of this protein family, in human breast carcinoma cells is critically involved in preventing EphA2 receptor degradation, and this activity is mediated by Odin Sam domains (Kim et al., 2009).

The EphA2 receptor is also involved in pathologies different than cancer. For example, EphA2 activation has been shown to mediate inflammation during injury, ischemia, and other chronic inflammatory conditions (Ivanov and Romanovsky, 2006; Kitamura et al., 2008) by promoting vascular permeability (Cercone et al., 2009); in particular, recently EphA2 has been related to exacerbation of ischemic brain injury (Thundyil et al., 2013).

A few studies have shown that mutations or deletions affecting the Sam domain of EphA2 receptor lead to cataract formation in humans and mice (Shiels et al., 2008). It has been proposed that these mutations influence protein stability through changes of solubility and folding efficiency, which may cause cellular disorganization and lens opacity (Park et al., 2012).

As described above, a few EphA2 pathological activities involve its Sam domain, which thus could be considered a novel target for the development of therapeutic compounds.

1.3. The PI3K effector protein Arap3

Arap3 (Arf GAP, Rho GAP, Ankyrin repeat and PH domains) is a dual GTPase activating protein for Arf6 and RhoA, containing five PH (Pleckstrin Homology) domains and a Sam domain (Krugmann et al., 2002) (**Figure 4**).

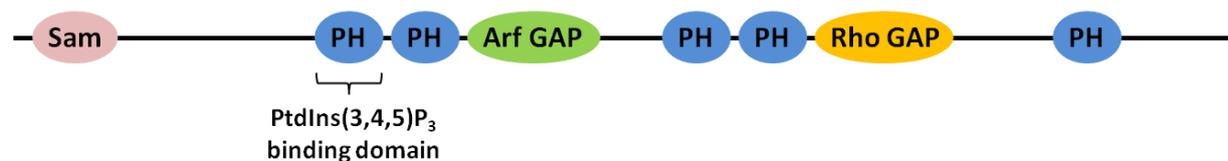


Figure 4. Domain organization of the protein Arap3.

Human Arap3 has been originally identified using an affinity-based method to purify proteins which specifically bind PtdIns(3,4,5)P₃ (Phosphatidylinositol-3-P) (Krugmann et al., 2002). PtdIns(3,4,5)P₃ is generated through phosphorylation of PtdIns(4,5)P₂ (Phosphatidylinositol-2-P) operated by the PI3K kinase (Vanhaesebroeck et al., 1997). Upon binding to PtdIns(3,4,5)P₃ with its N-terminal PH domain (Craig et al., 2010) (**Figure 4**), Arap3 translocates to the plasma membrane into proximity of its substrates, RhoA-GTP and Arf6-GTP, and activates them (Krugmann et al., 2004). Activation of these small GTPases is implicated in growth factors and PI3K kinase-

mediated signaling pathways, controlling lamellopodia formation, cell adhesion and spreading, regulation of actin cytoskeleton (I et al., 2004; Krugmann et al., 2006). Recently new Arap3 functions have been reported : it is a regulator of developmental sprouting angiogenesis (Gambardella et al., 2010), and a suppressor of peritoneal dissemination of scirrhous gastric carcinoma cells (Yagi et al., 2011).

In a yeast two-hybrid screen the lipid phosphatase Ship2 has been identified as binding partner for Arap3; this interaction is mediated by dimerization of Sam domains from the two proteins (Raaijmakers et al., 2007).

The NMR solution structure of the Arap3 Sam domain (Arap3-Sam) (PDB code: 2KG5) and its interaction with Ship2-Sam have been characterized (Leone et al., 2009; Raaijmakers et al., 2007). The Sam-Sam heterodimerization between proteins occur by means of the "Mid-Loop/End-Helix" topology (Leone et al., 2009) (See Paragraph 1.1), with a dissociation constant in the high nanomolar range (Raaijmakers et al., 2007). Arap3-Sam presents a sequence homology of about 58% with EphA2-Sam, and competes with it for binding to Ship2-Sam (Leone et al., 2009) (See Paragraph 1.2.1).

Ship2 is a negative regulator of the PI3K kinase pathway; since binding of Arap3 to PtdIns(3,4,5)P₃ is required for efficient membrane localization of the protein (Craig et al., 2010), dephosphorylation of PtdIns(3,4,5)P₃ by Ship2 implies a reduced affinity of Arap3 for the plasma membrane (Raaijmakers et al., 2007).

1.4. Odin: a member of the Anks family of proteins

The ANKS (Ankyrin repeat domain containing and Sam) protein family includes Odin and AIDA-1b (A β PP intracellular domain-associated protein 1B), both consisting of six ankyrin repeat domains at the N-terminus, two Sam domains in tandem, and a phosphotyrosine binding domain (PTB) at the C-terminus (Pandey et al., 2002; Ghersi et al., 2004) (**Figure 5**).



Figure 5. Domain arrangement of ANKS protein family.

Odin is a cytosolic adapter protein that negatively regulates PDGF (Platelet-Derived Growth Factor)-mediated cell proliferation (Pandey et al., 2002).

It has recently been reported a potential role of Sam domains of ANKS family proteins as inhibitors of EphA2 receptor degradation (Kim et al., 2010). In particular, Odin over-expression in MDA-MB-231 human breast carcinoma cells and MEFs (Mouse Embryonic Fibroblast) inhibits EphA2 internalization and degradation after ligand stimulation, whereas a Sam domain deletion mutant of Odin impairs this function (Kim et al., 2010). It has been suggested a fine regulation of EphA2 receptor signaling by a balance between the activity of c-Cbl E3 ubiquitin ligase (a positive regulator of Eph degradation) and ANKS proteins (Kim et al., 2010).

Sam domains of Odin have also been identified as binding partners of Arap3-Sam, but the possible function of these interactions remains unclear (Raaijmakers et al., 2007).

Before starting our studies the only structure information available in the PDB (Protein Data Bank) on Sam domains from the ANKS family concerned with AIDA-1b. In detail, the structures of the first Sam domain of AIDA-1 (PDB code: 2KE7-Donaldson, not published) and its Sam domain tandem (PDB code: 2KIV- Kurabi et al., 2010) had already been solved by NMR.

The AIDA-1b Sam domain tandem has a head-to-tail organization in which the Mid-Loop and End-Helix surfaces are provided by Sam1 and Sam2 respectively (**Figure 6**).



Figure6. NMR structure of the tandem Sam domain of AIDA-1 (first conformer of the NMR ensemble).

My research, as will be better describe below, focused on the first Sam domain of Odin.

1.5. Aims of the present PhD thesis

My thesis rely on an extensive analysis at molecular level of Sam domains-mediated protein-protein interactions, in order to lay the foundations for the design of libraries of biotechnological relevant molecules which, by mimicking the interaction surfaces and thus, selectively interfering with the formation of Sam-Sam complexes, could prove useful for either therapeutic and diagnostic applications.

In particular, I focused my studies on the solution structure and the intermolecular interactions of a few Sam domain containing proteins that play important roles in

physiological or pathological processes, and the analysis of solution conformational properties of peptides reproducing the corresponding binding sites.

This work is arranged into three main work packages:

1) The NMR solution structure of the first Sam domain of Odin and its interaction with the EphA2 receptor (Chapter 2).

The first Sam-Sam interaction I have characterized is the complex formation between the tyrosine kinase receptor EphA2 (Ephrin receptor A2), which is able to regulate important biological processes and play a fundamental role in tumorigenesis, and Odin, a protein belonging to the ANKS (Ankyrin repeat domain containing and Sam) family which possesses two Sam domains in tandem.

2) Studies of the heterotypic Sam-Sam association between Odin and Arap3 (Chapter 3).

Another Sam-Sam interaction which has been object of my studies is that occurring between Odin and Arap3 (Arf GAP, Rho GAP, Ankyrin repeat and PH domains), a protein involved in the phosphoinositol-3-kinase (PI3K) signaling pathways.

3) NMR structural characterization of the Sam3 peptide (Chapter 4).

A peptide, named Sam3, which includes the central portion of the Odin-Sam1 domain, that is mainly responsible, as demonstrated by my previous results, of the association with EphA2-Sam and Arap3-Sam, together with the $\alpha 5$ C-terminal helix, was designed (See Figure 1, Chapter 4). The C-terminal helix was added to induce the peptide folding in a similar way with respect to the intact Odin-Sam1 domain. I have analyzed the conformational properties of the peptide in buffer and H₂O/TFE (Trifluoroethanol) mixtures.

To achieve the above described goals, a multidisciplinary approach involving different experimental and computational techniques, such as NMR (Nuclear Magnetic Resonance Spectroscopy), SPR (Surface Plasmon Resonance), ITC (Isothermal Titration Calorimetry), docking, mutagenesis, was undertaken.

The knowledge acquired from these studies could be effectively the starting point for the design of biotechnological relevant molecules (like peptides and peptidomimetics) interfering with Odin-Sam1/EphA2-Sam or Odin-Sam1/Arap3-Sam associations. Particularly, antagonists of Odin-Sam1/EphA2-Sam interaction could modulate EphA2 receptor endocytosis and play a role as therapeutic agents in diseases such as cancer. With regard to the Odin-Sam1/Arap3-Sam complex, the designed molecules could be implemented in cellular assays to identify new biological functions related to Arap3-Sam.

2. THE NMR SOLUTION STRUCTURE OF THE FIRST SAM DOMAIN OF ODIN AND ITS INTERACTION WITH THE EPHA2 RECEPTOR

2.1. BACKGROUND

The EphA2 receptor belongs to a large subgroup of receptor tyrosine kinase family (Eph) playing relevant functions in several physiological and pathological events. In particular, EphA2 has long been related to cancer (Pasquale, 2010; Surawska et al., 2004), and recently has been associated to cataracts (Cooper et al., 2008; Jun et al., 2009; Zhang et al., 2009) and inflammation processes (Coulthard et al., 2012). These pathological functions of EphA2 have made it an attractive target in drug discovery (Koolpe et al., 2002; Noberini et al., 2008; Tagnolini et al., 2012).

With regard to EphA2 pro-cancer activities, the potential correlation between the process of receptor endocytosis and subsequent degradation, and the decrease of tumor malignancy, has been analyzed in many types of cancer in which EphA2 is over-expressed (Pasquale, 2010; Surawska et al., 2004). Recent studies have reported that the EphA2 endocytosis regulation is effected by some proteins, including the lipid phosphatase Ship2 (Zhuang et al., 2007). In fact, *in vitro* experiments have demonstrated that Ship2 over-expression in malignant breast cancer cells increases EphA2 stability, with a consequent enhancement in tumor cell invasiveness and metastatic potential, while decreased levels of Ship2 facilitate receptor internalization and degradation. This regulation process occurs as a result of an heterotypic interaction in between the Sam domain of Ship2 (Ship2-Sam) and the Sam domain of EphA2 (EphA2-Sam) (Zhuang et al., 2007).

The NMR (Nuclear Magnetic Resonance) solution structure of Ship2-Sam (PDB code: 2K4P) was previously reported, together with binding studies with the EphA2-Sam receptor (Leone et al., 2008). In particular, ITC (Isothermal Titration Calorimetry) studies were carried out and showed that EphA2-Sam binds Ship2-Sam with an affinity in the low micromolar range. Chemical shift perturbation studies have revealed the reciprocal binding interfaces of Ship2-Sam and EphA2-Sam, and shown that these proteins interact by adopting a topology common to several Sam-Sam complexes, called "Mid-Loop/End-Helix" (Leone et al., 2008).

A recent work indicates that the EphA2 receptor endocytosis is also regulated by the ANKS (Ankyrin repeat domain containing and Sam) protein family (Kim et al., 2010). This class of proteins includes Odin (Pandey et al., 2002) and AIDA1b (A β PP intracellular domain-associated protein 1B) (Gherzi et al., 2004), and is characterized by the presence of two Sam domains in tandem (Sam1 and Sam2). In particular, Odin over-expression in MDA-MB-231 human breast carcinoma cells and MEFs (Mouse Embryonic Fibroblast) inhibits EphA2 internalization and degradation after ligand stimulation, whereas a Sam domain deletion mutant of Odin impairs this function (Kim et al., 2010).

Within this work package, the solution structure of the first Sam domain of Odin (Odin-Sam1) was determined by means of NMR structural studies. Indeed, studies of Odin-Sam1/EphA2-Sam complex were carried out by means of NMR, SPR (Surface Plasmon Resonance), ITC and molecular docking techniques.

2.2. MATERIALS AND METHODS

2.2.1. Protein expression

Recombinant proteins were expressed in *Escherichia Coli*, using constructs provided with a tail of six His and a thrombin cleavage site at the N-terminus. Genes were cloned into the PET15B plasmids and transformed using BL21-Gold (DE3) competent cells (Stratagene). The following PET15B-constructs were used in this study: residues 1199-1258 of human Ship2 (UniprotKB/TrEMBL code: O15357), residues 691-770 of human Odin (UniprotKB/TrEMBL code: Q92625), residues 901-976 of human EphA2 (Swiss-Prot/TrEMBL code: P29317), EphA2-Sam double mutant (H924N, R950A corresponding to H45N, R71A according to our sequence number), and EphA2-Sam triple mutant (K917A, R957A, Y960S corresponding to K38A, R78A, Y81S according to our sequence number). Synthetic genes coding for these proteins were all purchased from Celtek Bioscience (Nashville, TN).

Unlabeled proteins expression was achieved by growing bacteria at 37°C in LB medium until $OD_{600}=0.6$. Protein over-expression was induced by IPTG (Isopropyl β -D-thiogalactopyranoside) at a concentration of 1 mM, overnight at 25°C. Expression of $^{15}\text{N}/^{13}\text{C}$ double labeled and ^{15}N uniformly labeled proteins was carried out in M9 minimal medium supported with 2 g/L of ^{13}C -Glucose and/or 0.5 g/L of $^{15}\text{NH}_4\text{Cl}$. M9 medium containing 3.6 g/L of ^{12}C -glucose (natural abundance) and 0.4 g/L of ^{13}C -glucose was implemented to achieve 10% fractional ^{13}C labeling for stereo-specific assignments of Leu- $\text{CH}_3^{\delta^{1,2}}$ /Val- $\text{CH}_3^{\gamma^{1,2}}$ methyl groups (Neri et al., 1989). The expression protocol for labeled proteins was identical to that used for unlabeled protein production. Proteins were purified with an AKTA prime FPLC system by affinity chromatography on a nickel column (Amersham).

2.2.2. Resonance assignment

Experiments for resonance assignments were acquired at 25 °C on a Varian Unity Inova 600 MHz spectrometer equipped with a cold probe. NMR samples were made up of ^{15}N or $^{15}\text{N}/^{13}\text{C}$ labeled Odin-Sam1 (900 μM) in phosphate buffer saline (PBS, 10 mM phosphates, 138 mM NaCl, 2.7 mM KCl) (Fisher) at pH=7.7, 0.2% NaN_3 with volumes of 600 μL (95% $\text{H}_2\text{O}/5\%$ D_2O). In **Figure 1** the [$^1\text{H},^{15}\text{N}$] HSQC (Heteronuclear Single Quantum Coherence) spectrum of ^{15}N labeled Odin-Sam1 is shown, and indicates that the protein is folded and not degraded.

Backbone assignments were carried out through analysis of triple resonance experiments (HNCA (**Figure 2A**), HN(CO)CACB, HNCACB) (Grzesiek and Bax, 1993). (H)CC(CO)NH and HCCH-TOCSY spectra were used to identify carbon side chains. Proton side chains were assigned analyzing the HCCH-TOCSY spectrum or by comparing 3D ^{15}N resolved- [$^1\text{H}, ^1\text{H}$] NOESY (100 ms mixing time) (**Figure 2B**) and 3D ^{15}N resolved- [$^1\text{H}, ^1\text{H}$] TOCSY (70 ms mixing time) experiments. 2D [$^1\text{H}, ^1\text{H}$] NOESY (mixing time 100 ms) and 2D TOCSY (mixing time 70 ms) experiments, recorded with samples of Odin-Sam1 dissolved in 99% D_2O , were recorded in order to obtain aromatic side chains assignments. Stereo-specific assignments for Leu- $\text{CH}_3^{\delta^{1,2}}$ and Val- $\text{CH}_3^{\gamma^{1,2}}$ methyl groups were obtained analyzing a [$^1\text{H}, ^{13}\text{C}$] HSQC experiment acquired on a fractionally ^{13}C labeled Odin-Sam1 sample (500 μM) (Neri et al., 1989).

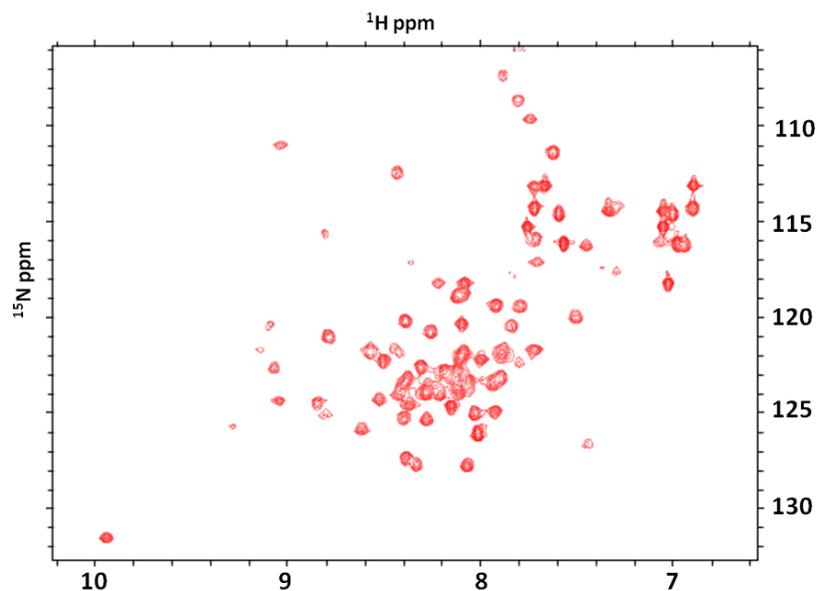


Figure 1. [^1H , ^{15}N] HSQC spectrum of ^{15}N labeled Odin-Sam1. Each peak corresponds to H_N backbone atoms and side chains H_N from Asn, Gln and Trp.

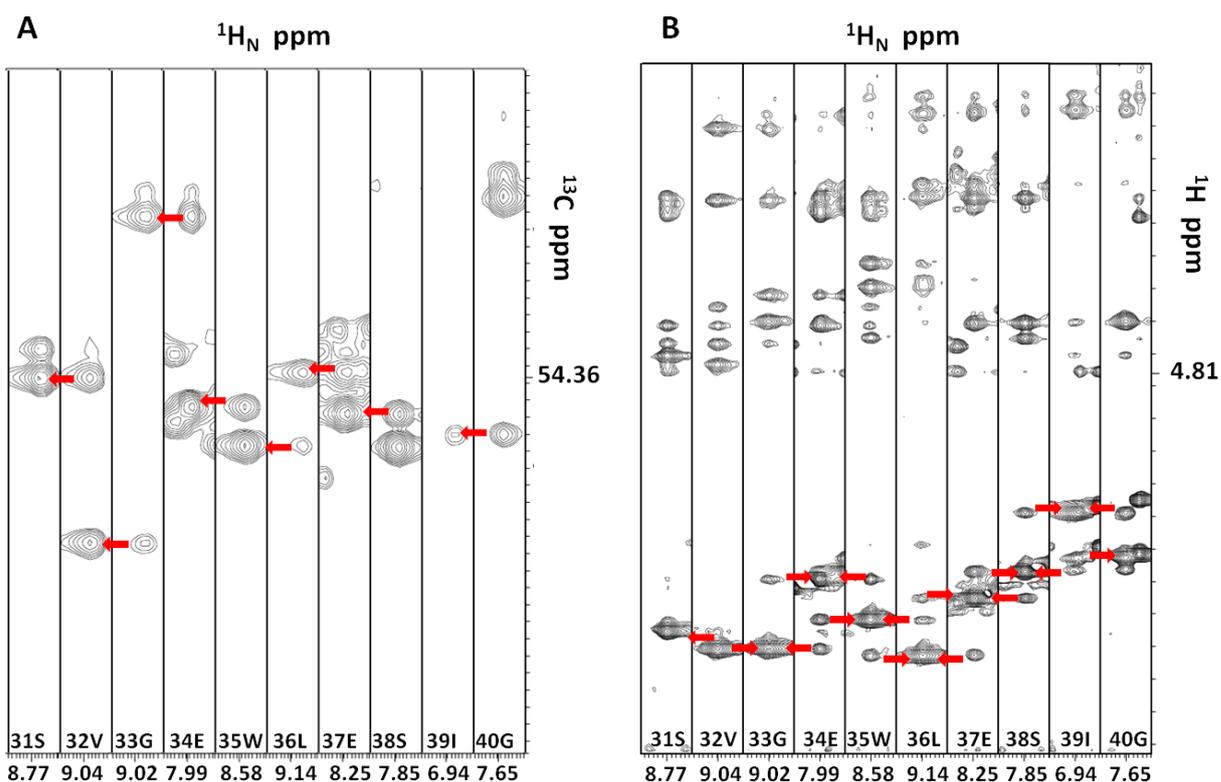


Figure 2. (A) A portion of the Odin-Sam1 HNCA spectrum encompassing residues from Ser31 to Gly40. Arrows indicate correlations between the C_α atoms of adjacent residues. (B) The region Ser31-Gly40 of Odin-Sam1 showed in its 3D ^{15}N resolved- $[\text{}^1\text{H}$, $^1\text{H}]$ NOESY spectrum. Arrows indicate sequential correlations between H_N backbone atoms.

To gain backbone assignments of the Odin-Sam1/EphA2-Sam complex, HNCA and 3D ^{15}N resolved- ^1H , ^1H] NOESY (100 ms mixing time) spectra were recorded on samples including either $^{15}\text{N}/^{13}\text{C}$ double labeled Odin-Sam1 (450 μM) and unlabeled EphA2-Sam (~1 mM), or double labeled EphA2-Sam (350 μM) and unlabeled Odin-Sam1 (900 μM).

NMR spectra were processed with Varian software (Vnmrj version 1.1D) and analyzed with NEASY as implemented in the program Cara (Bartels et al., 1995).

2.2.3. Relaxation measurements

The aggregation state of Odin-Sam1 in solution was analyzed by means of backbone ^{15}N R1 and R2 nuclear spin relaxation rates measurements at 25 °C and 600 MHz. These measurements were performed on ^{15}N uniformly labeled Odin-Sam1 samples at the concentrations of 100 μM and 900 μM , and on a sample containing uniformly ^{15}N labeled Odin-Sam1 (450 μM) and unlabeled EphA2-Sam (1 mM).

A series of 1D spectra (4K data points and 2K or 4K transients) were acquired to obtain R1 and R2 relaxation data: five relaxation delays (0.01, 0.1, 0.3, 0.6, 1.0 s) were used for R1 measurements and seven relaxation delays were implemented for R2 data sets (0.01, 0.03, 0.05, 0.07, 0.09, 0.11, 0.15, 0.19 s). The decrease in signal intensity as function of relaxation delay allowed to estimate average R1 and R2 values; then average R1/R2 ratios were used as input for the software tmest (A. G. Palmer III, Columbia University) (Kay et al., 1989) in order to calculate the rotational correlation time (τ_c).

2.2.4. DOSY (Diffusion Ordered Spectroscopy) experiments

The diffusion-ordered NMR spectroscopy was carried out with the Pulsed Gradient Spin-Echo (PGSE) NMR technique (Morris and Johnson, 1992). The following relationship correlates the translational self-diffusion coefficient, D , to the NMR parameters : $I=I_0 \exp[-(D\gamma^2\delta^2G^2(\Delta-\delta/3))]$, where I_0 is the measured peaks intensity of a particular group of resonances at the smaller gradient strength, I is the corresponding observed peak intensity, D is the translational self-diffusion constant, γ is the proton gyromagnetic ratio, δ is the diffusion gradient duration, G is the gradient strength, and Δ is the time gap between two gradients (Price et al., 1997).

Series of spectra were acquired with 512 scans and 16 K data points. DOSY experiments of Odin-Sam1 samples at the concentrations of 100 μM and 900 μM were recorded. The hydrodynamic radius r_H of the protein was evaluated with the Stokes-Einstein equation $D=k_B T/f$, where $f=6\pi\eta r_H$ is the translational friction coefficient, η is the viscosity of the solution, k_B is the Boltzmann constant, T is the temperature in K.

2.2.5. Structure calculation and analysis of Odin-Sam1

In order to collect distance constraints for structure calculations, a 3D ^{15}N resolved- ^1H , ^1H] NOESY-HSQC spectrum (Talluri et al., 1996) (100 ms mixing time), a 3D ^{13}C resolved- ^1H , ^1H] NOESY-HSQC spectrum (100 ms mixing time) and a 2D ^1H , ^1H] NOESY spectrum (Kumar et al., 1980) (100 ms mixing time) for the aliphatic to aromatic region, acquired after dissolving the lyophilized protein sample in 99% D_2O , were used. The software CYANA version 2.1 (Herrmann et al., 2002) was employed to calculate the solution structure of Odin-Sam1, and angular constraints were

generated with the GRIDSEARCH module of CYANA. Structure calculations were initiated from 100 random conformers; the 20 structures with the lowest CYANA target functions were analyzed with the programs MOLMOL (Koradi et al., 1996) and validated with iCING (<http://proteins.dyndns.org/cing/iCing.html>).

2.2.6. Odin-Sam1/EphA2-Sam binding studies

For chemical shift perturbation studies (Pellecchia, 2005) 2D [^1H , ^{15}N] HSQC experiments were recorded; residues close to the binding site can be identified by monitoring changes in the HSQC spectra of a protein upon addition of a ligand. The addition of ligand is stopped only when the saturation point is reached (i.e.: at this point any further addition of the binding partner does not produce other changes in the spectra).

Thus, to map the Odin-Sam1 binding interface for EphA2-Sam, 2D [^1H , ^{15}N] HSQC spectra of ^{15}N labeled Odin-Sam1 (160 μM) were acquired in presence and in absence of unlabeled EphA2-Sam (80 μM , 160 μM , 240 μM , 400 μM). To recognize the binding site of EphA2-Sam for Odin-Sam1, 2D [^1H , ^{15}N] HSQC spectra of a ^{15}N labeled EphA2-Sam sample (63 μM) were recorded in absence and presence of unlabeled Odin-Sam1 (360 μM and 600 μM). Chemical shift perturbation experiments were also performed to verify the binding of EphA2-Sam double (H45N, R71A) and triple (K38A, R78A, Y81S) mutants to Odin-Sam1. For this purpose, ^{15}N labeled Odin-Sam1 (50-100 μM) and unlabeled double and triple EphA2-Sam mutants (concentrations ranging from 200 μM to 2 mM) samples were used.

Furthermore, displacement experiments were carried out, in which a 2D [^1H , ^{15}N] HSQC spectrum of ^{15}N labeled Odin-Sam1 (150 μM) in its unbound state, after addition of unlabeled EphA2-Sam (363 μM), and next further addition of unlabeled Ship2-Sam (ratio Odin-Sam1/EphA2-Sam/Ship2-Sam equal to 1/2.4/13) were recorded.

Analysis of titration experiments and overlays of 2D spectra were generated with the program Sparky (Goddard and Kneller, SPARKY 3, University of California, San Francisco).

2.2.7. SPR experiments

SPR experiments were carried out at 25 °C and a constant 20 $\mu\text{L}/\text{min}$ flow rate, using as running buffer a solution of Hepes 10 mM, pH=7.4, NaCl 150 mM, surfactant P20 0.05% v/v (90 μL injected for each experiment). EphA2-Sam was immobilized in 10 mM acetate buffer pH 5.0 (flow rate 5 $\mu\text{L}/\text{min}$, time injection 7 min) on a CM5 Biacore sensor chip, using EDC/NHS chemistry (Johnsson et al., 1991), while Odin-Sam1 was used as an analyte at different concentrations in the range of 0.1-400 μM . Residual reactive groups were deactivated with 1 M ethanolamine hydrochloride, pH=8.5; the reference channel was prepared by activating with EDC/NHS and deactivating with ethanolamine. Immobilization levels were: 940, 1313, 1480 RU for EphA2-Sam wild-type, double and triple mutants, respectively. The BIA evaluation analysis package (version 4.1, GE Healthcare, Milano, Italy) was used to subtract the signal of the reference channel and to estimate K_D (Dissociation constant) values. RU_{max} (Resonance Units) values of each experiment versus Odin-Sam1 concentration were fit by non-linear regression analysis with GraphPad Prism, version 4.00 (GraphPad Software, San Diego, California) (Rich et al., 2001).

2.2.8. ITC experiments

ITC studies were performed at 25 °C with an iTC200 calorimeter (MicroCal/GE Healthcare, Milano, Italy). A solution of EphA2-Sam at a concentration of 257 μ M was titrated into a solution of Odin-Sam1 10 μ M concentrated. Before ITC measurements both proteins were widely dialyzed in the same buffer (PBS, pH 7.7). Fitting of data to a single binding site model was carried out with the Origin software as supplied by GE HealthCare.

2.2.9. Docking studies

Docking studies were carried out by using the NMR structures (first conformers) of both Odin-Sam1 (PDB code: 2LRM (Mercurio et al., 2012)) and EphA2-Sam (PDB code: 2E8N, Riken Structural Genomic Initiative). Models of the Odin-Sam1/EphA2-Sam complex were generated with the Haddock web server (de Vries et al., 2010). Ambiguous interaction restraints were generated from chemical shift perturbation data: residues of Odin-Sam1 and EphA2-Sam with the highest chemical shift variations and large solvent exposure, or which could potentially supply important intermolecular contacts, as revealed by structural homologies with other Sam-Sam complexes, were set as active. For Odin-Sam1 active residues are: L49, L50, N51, F53, D54, D55, V56, H57, F58, Q67, D68, R70, D71. For EphA2-Sam residues K38, R71, G74, H75, K77, R78, Y81 were set up as active. The regions encompassing residues 51-65 of Odin-Sam1 and residues 74-81 of EphA2-Sam (corresponding to amino acids 59-66 according to the sequence numbers of the PDB file 2E8N) were set as semi-flexible. The C-terminal and N-terminal tails of both Odin-Sam1 and EphA2-Sam were considered fully flexible during all the docking stages. In all the docking runs, passive residues were set automatically by the Haddock web server and the solvated docking mode was used (van Dijk and Bonvin, 2006).

The docking protocol included a first step (i.e.: the rigid body energy minimization) in which 1000 structures were calculated, in the second step the best 200 solutions were subjected to semi-flexible simulated annealing, then a final refinement in water was performed.

With all 200 final Haddock models a visual screening was carried out, with MOLMOL (Koradi et al., 1996), and solutions not compatible with our experimental data and/or containing atypical protein orientations were discarded. Seventy-eight structures were selected and further analyzed and compared with experimental structures of other heterotypic Sam-Sam complexes (such as the AIDA-1b Sam tandem, PDB code: 2KIV (Kurabi et al., 2009)), to identify common features. Finally, 22 models representing the possible EphA2-Sam/Odin-Sam1 complex conformations were chosen. The selected ensemble was clustered using a RMSD (Root Mean Square Deviation) cutoff value of 2 Å by using MOLMOL (Koradi et al., 1996). This clustering procedure was performed by superimposing the structures on the backbone atoms of the putative binding interfaces of EphA2-Sam (regions 36-41, 62-65, 70-81) and Odin-Sam1 (region 48-73), indicating the presence of 5 families of structures (See Appendix 1).

2.3. RESULTS AND DISCUSSION

EphA2 receptor tyrosine kinase is considered a promising target in drug discovery for cancer therapies; in fact, it is over-expressed in many aggressive tumors and exhibits several pro-malignancy activities. In particular the process of receptor endocytosis has been exploited as a possible route to reduce tumor malignancy (Pasquale, 2010; Surawska et al., 2004). Recent evidences have shown that Sam domains, small protein binding modules made up of a five helix bundle, are crucial for anchorage of protein regulators of endocytosis at the receptor site (Kim et al., 2010; Zhuang et al., 2007). The lipid phosphatase Ship2 is engaged in a heterotypic Sam-Sam interaction with the receptor and is able to inhibit its endocytosis in cancer cells (Zhuang et al., 2007). Sam domains of proteins belonging to the ANKS family also play an important role in the process, possibly by regulating ubiquitination mechanisms (Kim et al., 2010). Among ANKS family members, herein attention was focused on Odin (Pandey et al., 2002), which in cancer cells inhibits EphA2 endocytosis, whereas its mutant, lacking Sam domains, does not (Zhuang et al., 2010). Odin contains two Sam domains in tandem at the C-terminal side, Odin-Sam1 and Odin-Sam2. For Odin-Sam1 I have carried out a complete structural characterization by NMR. Moreover, due to the high sequence identity between Ship2-Sam and Odin-Sam1, as well as the common function of regulators of EphA2 endocytosis, I have investigated if Odin-Sam1 could directly bind EphA2-Sam as Ship2-Sam does.

2.3.1. Analysis of the aggregation state of Odin-Sam1

Before starting a complete structure calculation, I have investigated the aggregation state of Odin-Sam1 in solution by means of ^{15}N R1 and R2 nuclear spin relaxation rates measurements and DOSY experiments.

Odin-Sam1 has a rotational correlation time (τ_c), estimated by the R2/R1 average value, of 7.4 ± 0.7 ns at a protein concentration of $100 \mu\text{M}$, which does not increase at a protein concentration of $900 \mu\text{M}$ (7.3 ± 0.6 ns). These values are rather close to those reported for monomeric Sam domains such as Ship2-Sam (6.7 ns) (Leone et al., 2008) and Arap3 (Arf GAP, Rho GAP, Ankyrin repeat and PH domain)-Sam (8.2 ns) (Leone et al., 2009) at similar concentrations and buffer conditions.

The τ_c of Odin-Sam1 bound to EphA2-Sam increases instead to 11 ± 1 ns, reflecting the raise in molecular weight upon association; moreover this value is close to the τ_c measured for Ship2-Sam/EphA2-Sam (11.2 ± 0.5 ns) (Leone et al., 2008) and Ship2-Sam/Arap3-Sam (11 ± 1 ns) (Leone et al., 2009) complexes, thus suggesting that Odin-Sam1 and EphA2-Sam are forming a complex with a 1:1 stoichiometry.

By means of DOSY experiments I have measured a diffusion coefficient (D) for the diluted Odin-Sam1 sample ($100 \mu\text{M}$) of $1.46 \cdot 10^{-10} \text{ m}^2 \text{ s}^{-1}$, corresponding to a hydrodynamic radius (r_H) value of 16.7 \AA , and a D for the more concentrated Odin-Sam1 sample ($900 \mu\text{M}$) equal to $1.40 \cdot 10^{-10} \text{ m}^2 \text{ s}^{-1}$, corresponding to a r_H of 17.5 \AA . DOSY measurements show that the hydrodynamic radius of the protein (i.e.: $\sim 17 \text{ \AA}$) results comparable with that of compact proteins of similar size (Wilkins et al., 1999). The results obtained from relaxation and DOSY studies indicate that aggregation phenomena can be excluded under the conditions that have been used to calculate the NMR structure (i.e.: $900 \mu\text{M}$) and confirm the Sam domains weak tendency to associate in solution through homotypic interactions (Thanos et al., 1999).

2.3.2. NMR solution structure of Odin-Sam1

To identify closely related Sam domains, a *blastp* search against the PDB database (Altschul et al., 1997) was carried out by using Odin-Sam1 sequence as input query. Odin-Sam1 (UniprotKB/TrEMBL code: Q92625, residues 696-762) presents with AIDA1b-Sam1 (UniprotKB/TrEMBL code: Q7Z6G8, residues 810-876) the maximum sequence identity (57%). Odin-Sam1 has also good sequence identity with Ship2-Sam (48%, UniprotKB/TrEMBL code: O15357, residues 1196-1258). A similar *blastp* search conducted by using Odin-Sam2 sequence (UniprotKB/TrEMBL code: Q92625, residues 770-837) indicates highest identity with AIDA1b-Sam2 (59%, UniprotKB/TrEMBL code: Q7Z6G8, residues 884-949) and 38% of sequence identity with EphA2-Sam (Swiss-Prot/TrEMBL: P29317, residues 904- 971) (**Figure 3**).

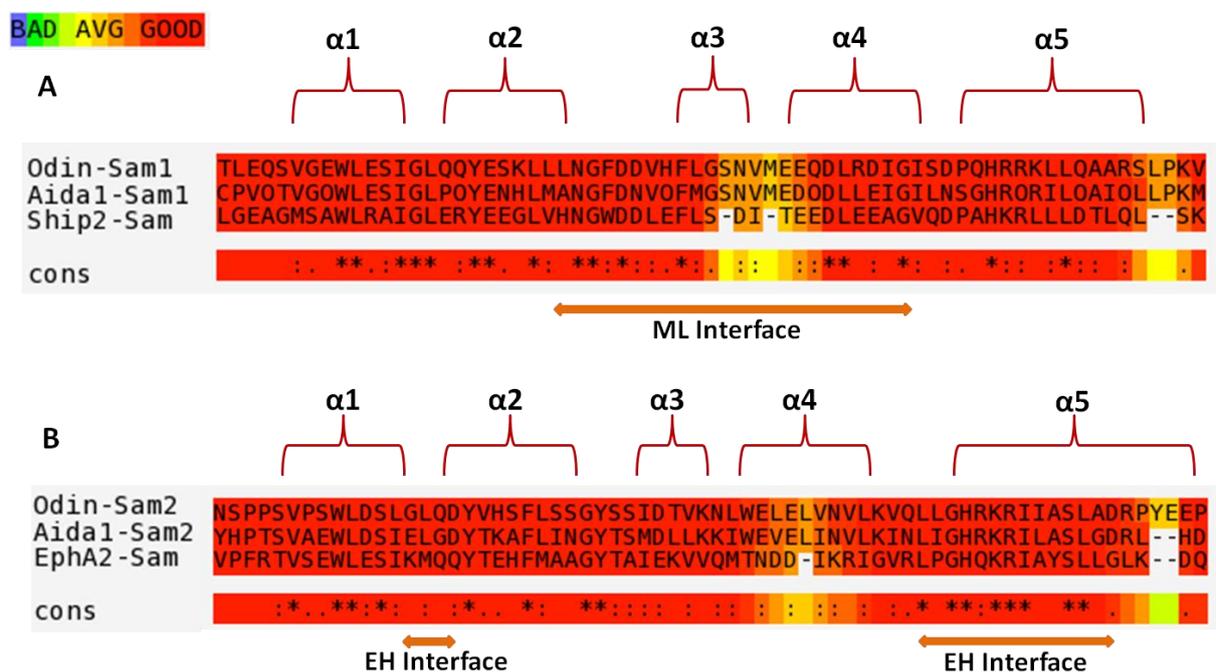


Figure 3. (A) Sequence alignment of human Odin-Sam1 with AIDA1b-Sam1 and Ship2-Sam. Secondary structure elements of Odin-Sam1 and residues contributing to ML surface are indicated. (B) Sequence alignment of Odin-Sam2 with AIDA1b-Sam2 and EphA2-Sam. Secondary structure elements and regions corresponding to the EH interface of EphA2-Sam are indicated with parenthesis and arrows respectively. Alignments have been generated with the T-COFFEE server (Notredame et al., 2000). Red residues correspond to highly reliable portions of the multiple alignment. The Cons line is a consensus, it indicates the average reliability value for each column.

The 3D solution structure of Odin-Sam1 was calculated with the program CYANA (Herrmann et al., 2002), which performs simulated annealing with torsion angle dynamics by using distance constraints from NOESY experiments. Odin-Sam1 solution structure consists of 5 α helices (**Figure 4A**), and is a canonical Sam domain helix bundle, in agreement with the high homologies with other Sam domains revealed by *blastp* (**Figure 3**). The final CYANA structure calculation includes 1206 upper distance constraints (393 intra-residue, 239 short-range, 275 medium-range, 299 long-range), 372 angle constraints (**Figure 4C**), and information on

stereospecific assignments for methyl groups of Val32, Val56, Val63, Leu36, Leu48 and Leu84.

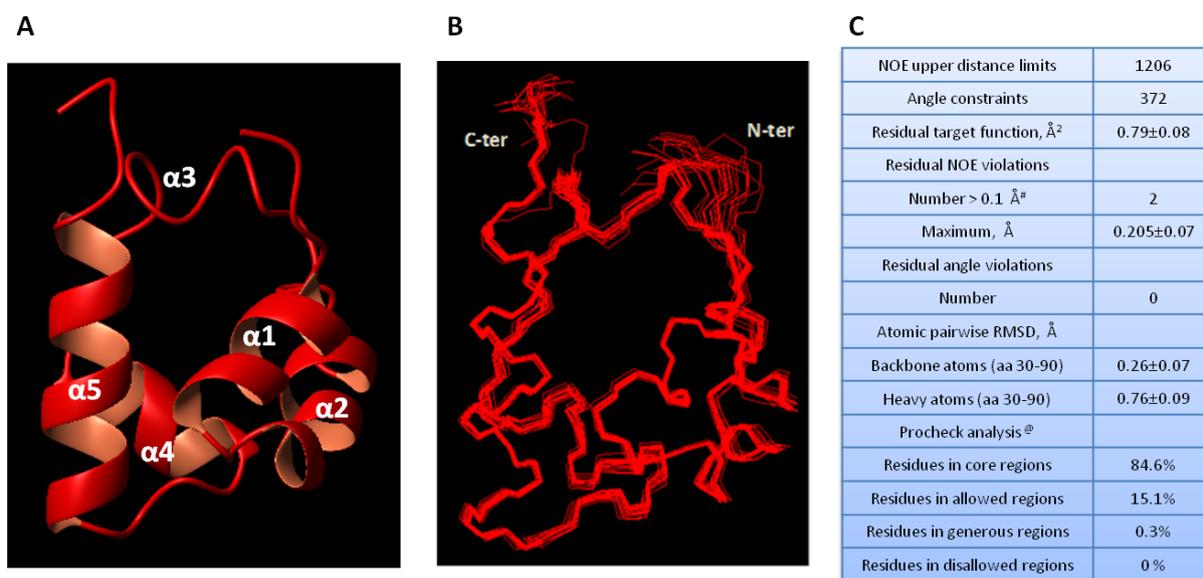


Figure 4. (A) Ribbon representation of Odin-Sam1 solution structure (first conformer). It contains the following α helical segments: $\alpha 1$ (residues 32-39), $\alpha 2$ (residues 42-50), $\alpha 4$ (residues 66-72), $\alpha 5$ (residues 77-88) (B) Superposition on the backbone atoms (region 30-90) of Odin-Sam1 NMR structures. (C) Structure features of Odin-Sam1 solution structure: CYANA (Herrmann et al., 2002) and PROCHECK_NMR (Laskowski et al., 1996) statistics for residues 30-90.

2.3.3. Odin-Sam1/EphA2-Sam interaction studies

The interaction between Odin-Sam1 and EphA2-Sam was investigated by means of NMR, SPR and ITC studies. A clear association of the two proteins is evident from analysis of all the different binding assays.

To map the Odin-Sam1 binding interface for EphA2-Sam, 2D [¹H, ¹⁵N] HSQC spectra of ¹⁵N labeled Odin-Sam1 were acquired in presence and in absence of unlabeled EphA2-Sam. Average ¹H and ¹⁵N chemical shifts variations were estimated with the equation $\Delta\delta = [(\Delta H_N)^2 + (0.17 \times \Delta^{15}N)^2]^{1/2}$ (Farmer et al., 1996). Amino acids presenting largest $\Delta\delta$ are expected to contribute to the binding surface. In particular, largest $\Delta\delta$ (values > 0.2 ppm) were mainly found in the middle part of the protein, including the C-terminal portion of $\alpha 2$ helix, the $\alpha 3$ helix and the N-terminal portion of the $\alpha 4$ helix (Figure 5).

Similar titration experiments were acquired to identify the binding interface of EphA2-Sam for Odin-Sam1, by using ¹⁵N labeled EphA2-Sam and unlabeled Odin-Sam1. Residues affected by the interaction were assessed by analyzing the normalized chemical shift deviations. The largest variations were observed for residues belonging to the $\alpha 5$ helix and the adjacent $\alpha 2\alpha 3$ and $\alpha 4\alpha 5$ loop regions (Figure 6).

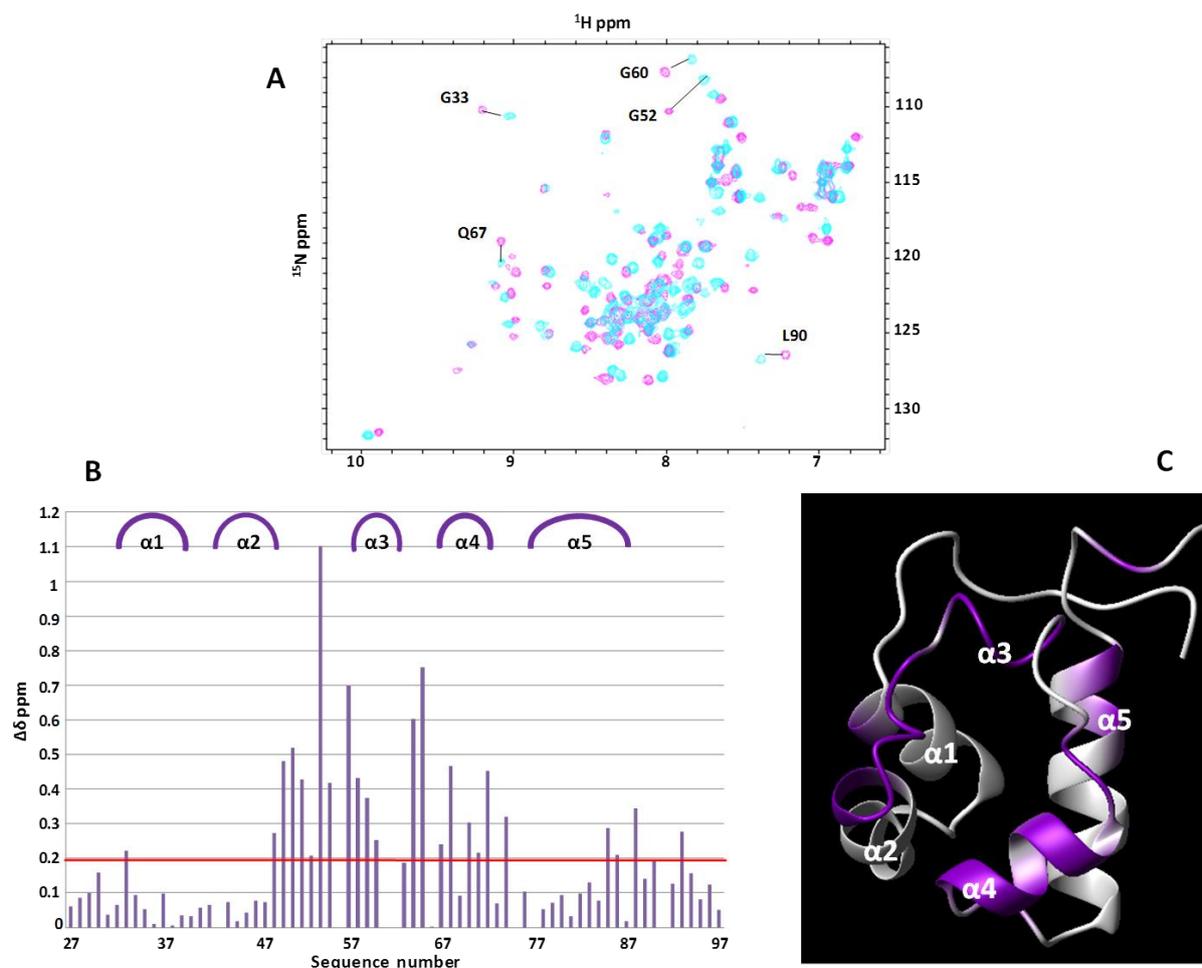


Figure 5: (A) Comparison of [^1H , ^{15}N] HSQC spectra of Odin-Sam1 (150 μM) in its unbound form (cyan) and after addition of EphA2-Sam (363 μM) (magenta). (B) Histogram showing normalized chemical shift deviations as function of the residue number. Residues G33, L49, L50, N51, G52, F53, D54, D55, H57, F58, L59, G60, M64, E65, Q67, D68, R70, D71, I72, I74, Q85, R88, V93 present normalized deviations with values higher than 0.2 ppm. $\Delta\delta$ values equal to 0 have been assigned to residues Q43, V56, N62 (their peaks only appear in the spectrum of the complex), S61 and S75 (unassigned), P77 and P91. (C) Residues with normalized chemical shifts deviation higher than 0.2 ppm are represented in dark violet on the 3D solution structure of Odin-Sam1 (PDB code: 2LMR (Mercurio et al., 2012)) in its ribbon representation.

These data seem to indicate that Odin-Sam1 and EphA2-Sam adopt for the heterotypic Sam-Sam interaction a Mid-Loop/End Helix model that is characteristic of Sam-Sam complexes (Kurabi et al., 2009; Rajakulendran et al., 2008; Ramachander et al., 2004; Stafford et al., 2011), in which Odin-Sam1 provides the Mid-Loop interface and EphA2-Sam provides the End-Helix interface.

SPR and ITC experiments were also performed, and both confirmed a clear association of Odin-Sam1 and EphA2-Sam (**Figure 7**).

For SPR experiments, EphA2-Sam was immobilized on the chip surface while Odin-Sam1 was used as an analyte; kinetic experiments in which R_Umax values of each experiment were plotted *versus* Odin-Sam1 concentration, both employing a 1:1 interaction model, provided a dissociation constant (K_D) for the Odin-Sam1/EphA2-

Sam complex (**Figure 7A, B**) of $5.5 \pm 0.9 \mu\text{M}$, obtained by best fitting of experimental data with a non-linear regression analysis (**Figure 7B**).

ITC experiments indicated instead that Odin-Sam1 associated with EphA2-Sam with a 1:1 stoichiometry, and a K_D equal to $0.62 \pm 0.04 \mu\text{M}$ (**Figure 7C**).

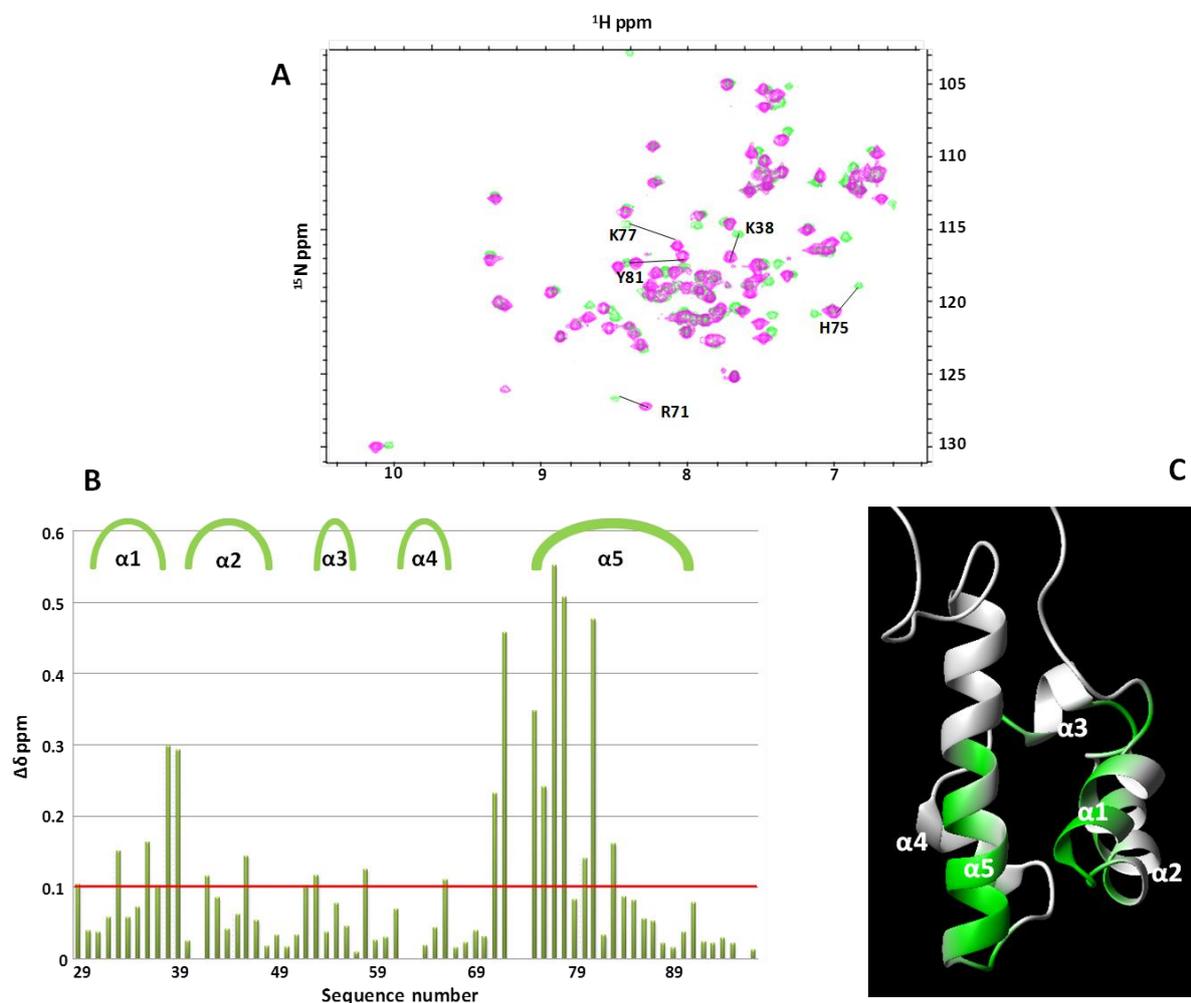


Figure 6: (A) Comparison of [^1H , ^{15}N] HSQC spectra of EphA2-Sam ($63 \mu\text{M}$) in its unbound form (magenta) and after addition of Odin-Sam1 ($360 \mu\text{M}$) (green). (B) Histogram showing normalized chemical shift deviations as function of the residue number. T29, W33, S36, I37, K38, M39, Y42, F46, T52, A53, V58, K66, R71, L72, H75, Q76, K77, R78, A80, Y81, L83 present normalized deviations with values higher than 0.1 ppm. $\Delta\delta$ values equal to 0 have been assigned to residues N62 and D63 (unassigned), G74, Q41 (their peaks only appear in the spectrum of the complex), P73 and P96. (C) Residues with normalized chemical shifts deviations higher than 0.1 ppm are represented in green on the 3D solution structure of EphA2 (PDB code: 2E8N, RIKEN Structural Genomics Initiative) in its ribbon representation.

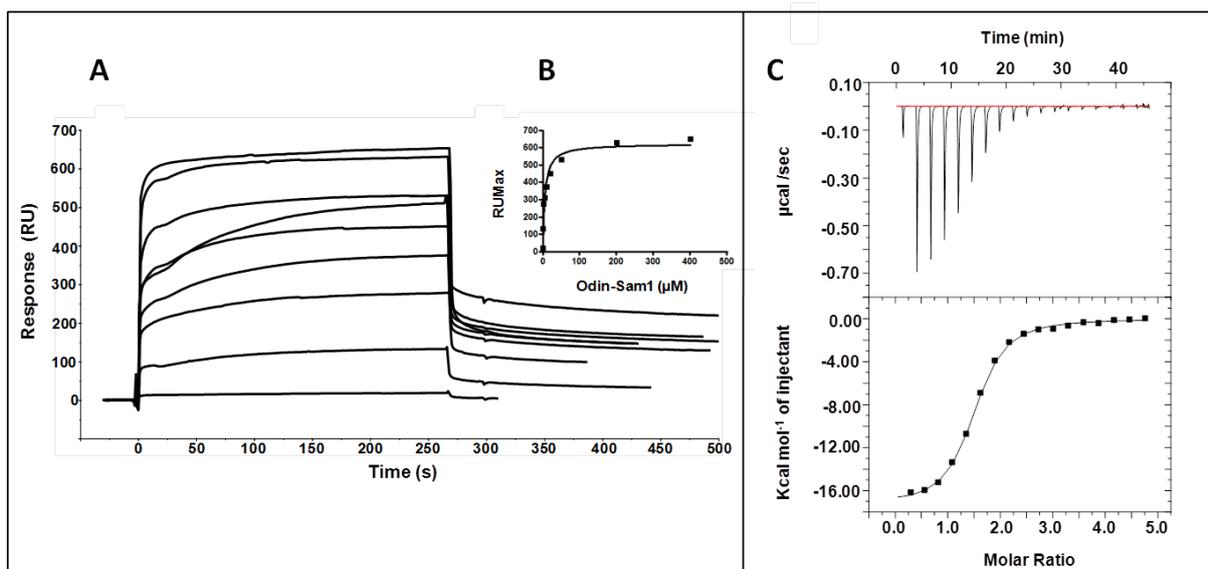


Figure 7. (A) SPR studies: overlay of sensorgrams representing the direct binding of Odin-Sam1 to immobilized EphA2-Sam (0.1-400 μM) (B) A plot of RU_{max} from each binding *versus* Odin-Sam1 concentration. (C) ITC studies: calorimetric curve relative to the titration of EphA2-Sam (257 μM) with Odin-Sam1 (10 μM) are showed. The row and the integrated data are reported in the top and bottom sections respectively. For data fitting (bottom) a single-binding model was applied (Mercurio et al., 2012).

Additional information on the binding mode of EphA2-Sam to Odin-Sam1 were derived by further NMR and SPR experiments with a triple mutant (K38A, R78A, and Y81S) and a double mutant (H45N and R71A) of EphA2-Sam.

In the triple EphA2-Sam mutant residues located in the $\alpha 5$ helix and $\alpha 1\alpha 2$ loop, that constitute the putative interaction surface of EphA2-Sam for Odin-Sam1 according to the results of chemical shift perturbation experiments, were mutated (**Figure 8**). The type of amino acids replacements were planned to destroy potential key interactions at the dimer interface without perturbing the overall protein structure. Instead, in the H45N, R71A EphA2-Sam mutant, amino acids replacements were inserted in regions adjacent to the putative binding site (**Figure 8**). Both NMR (**Figure 9A, B**) and SPR experiments (**Figure 9C, D**) indicate that, compared with EphA2-Sam wild-type, the binding affinity of the triple mutant for Odin-Sam1 is lower, whereas EphA2-Sam double mutant preserves similar ability to interact with Odin as the wild-type protein ($K_D = 3.8 \pm 0.3 \mu\text{M}$).

To recognize if EphA2-Sam was interacting with Odin-Sam1 and Ship2-Sam by using the same binding pocket, displacement experiments were carried out by means of 2D [^1H , ^{15}N] HSQC spectra. A spectrum of ^{15}N labeled Odin-Sam1 in its unbound state was first recorded. After addition of unlabeled EphA2-Sam to this sample, changes in the 2D [^1H , ^{15}N] HSQC characteristic of the Odin-Sam1/EphA2-Sam complex could be observed. The spectrum of unbound form of Odin-Sam1 could be restored by further addition of unlabeled Ship2-Sam (ratio Odin-Sam1/EphA2-Sam/Ship2-Sam equal to 1/2.4/13) (**Figure 10**).

These data highlight that the dissociation of the Odin-Sam1/EphA2-Sam complex is followed by capture of EphA2-Sam by Ship2-Sam and thus show that EphA2-Sam presents one identical binding pocket for Odin-Sam1 and Ship2-Sam.

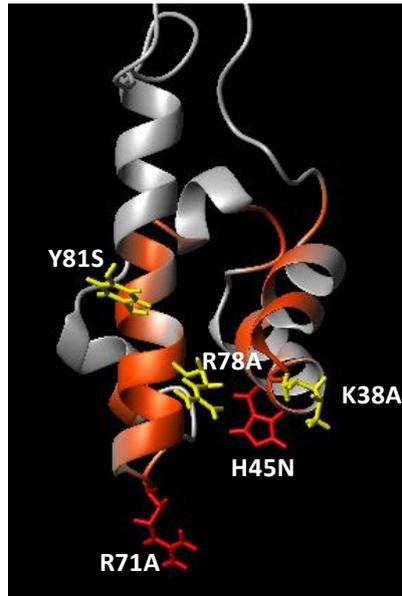


Figure 8. Ribbon representation of EphA2-Sam NMR solution structure (first conformer, PDB code: 2E8N), with residues corresponding to the binding interface for Odin-Sam1 highlighted in orange. The side chain of amino acids replaced in EphA2-Sam triple and double mutants are shown in neon representation (yellow and red respectively).

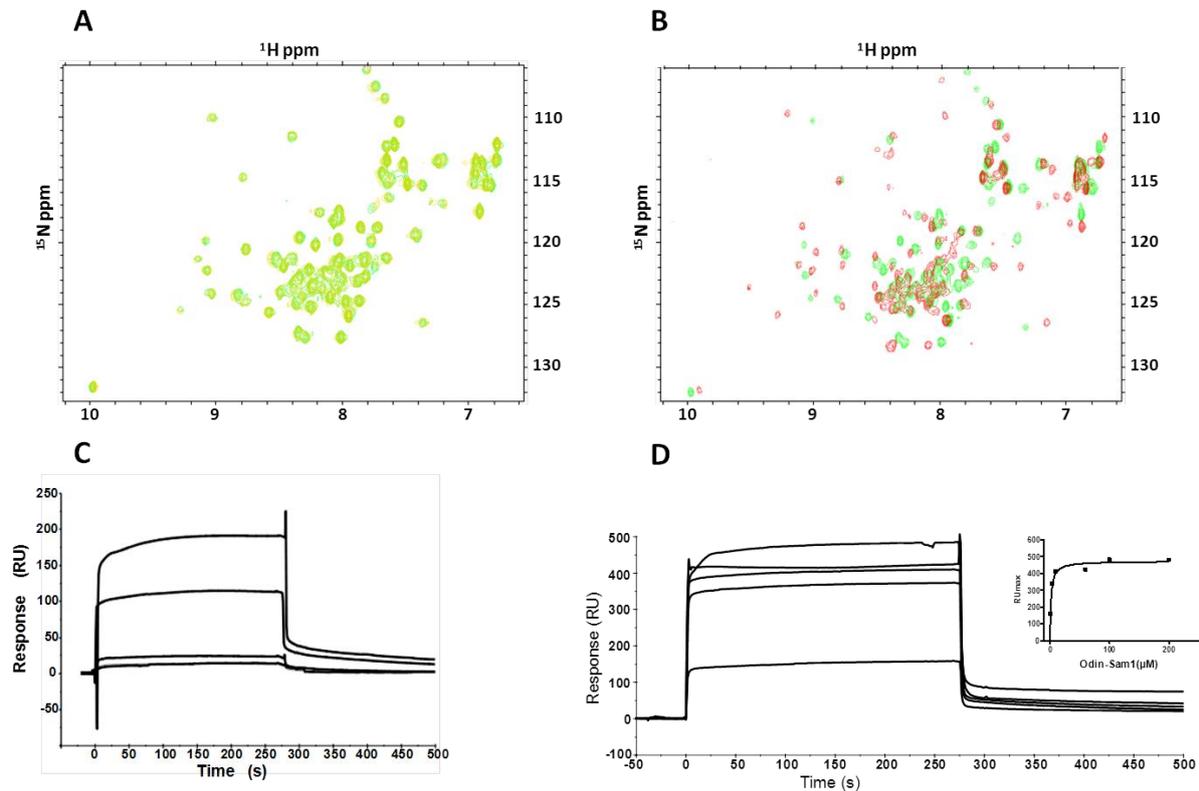


Figure 9. (A) Overlay of 2D [^1H , ^{15}N] HSQC spectra of ^{15}N labeled Odin-Sam1 in absence (green) and in presence (yellow) of EphA2-Sam triple mutant. (B) 2D [^1H , ^{15}N] HSQC spectra of ^{15}N labeled Odin-Sam1 in its unbound form (green) and after addition of EphA2-Sam double mutant (red). (C) Overlay of sensorgrams relative to the direct binding of Odin-Sam1 to immobilized EphA2-Sam triple mutant (1-200 μM). Saturation was not reached (Mercurio et al., 2012). (D) Overlay of sensorgrams relative to the direct binding of Odin-Sam1 to immobilized EphA2-Sam double mutant (1-200 μM). In the right inset a plot of RU_{max} from each binding versus Odin-Sam1 concentrations is reported; data were fitted by non-linear regression analysis (Mercurio et al., 2012).

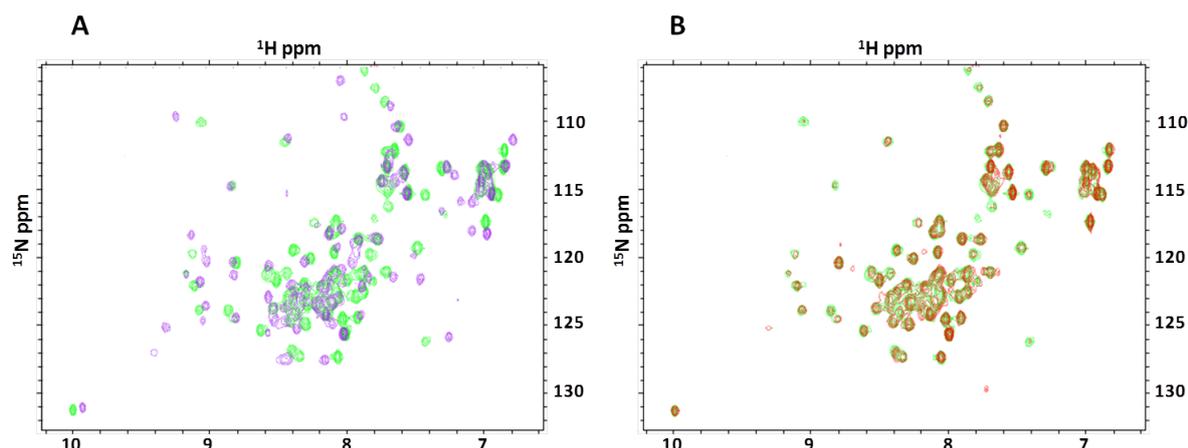


Figure 10. (A) Overlay of 2D [^1H , ^{15}N] HSQC spectra of ^{15}N labeled Odin-Sam1 (150 μM) in the apo form (green) and bound to EphA2-Sam (363 μM) (purple). (B) The spectrum of unbound Odin-Sam1 is restored by addition of unlabeled Ship2-Sam (red).

2.3.4. Molecular docking of the Odin-Sam1/EphA2-Sam complex

A model of the Odin-Sam1/EphA2-Sam complex was built by molecular docking with the Haddock web server (de Vries et al., 2010), by using the 3D structures of both Odin-Sam1 (PDB code: 2LMR) and EphA2-Sam (PDB code: 2E8N).

These modeling studies, based on chemical shift perturbation data, indicate that the Odin-Sam1/EphA2-Sam complex possesses a head-to-tail topology, also called Mid-Loop (ML)/End-Helix (EH) (Kim et al., 2002), in which Odin-Sam1 and EphA2-Sam provides the ML (i.e.: central region of the Sam domain) and EH (i.e.: C-terminal $\alpha 5$ helix and adjacent loops) binding interfaces, respectively. Mainly intermolecular interactions consist of H-bonds and electrostatic contacts in between positively charged residues of EphA2-Sam and negatively charged residues of Odin-Sam1 (**Figure 11**). Cation- π and π - π interactions also occur in a few Haddock solutions, and involve Phe53 and Phe58 on Odin-Sam1 surface and Tyr81 on the EphA2-Sam EH site (**Figure 11**).

Mutagenesis studies indicate indeed that Tyr81, together with Lys38 and Arg77 on the EH surface of EphA2 are likely providing essential interactions, in fact, concurrent mutations of the Tyr to Ser, Lys and Arg to Ala, highly attenuate the binding affinity for Odin-Sam1 (**Figure 9A, C**). Indeed, in most of our docking solutions these three residues are involved in intermolecular contacts (See Appendix 1).

As previously reported, Tyr81, Lys38 and Arg78 are also important for the interaction of EphA2-Sam with the Sam domain of the lipid phosphatase Ship2 (Ship2-Sam) (Leone et al., 2008). These earlier studies have pointed out that Ship2-Sam and EphA2-Sam bind to each other with a $K_D=0.75$ μM , probably adopting a ML/EH binding topology, in which Ship2-Sam and EphA2-Sam are providing the ML and EH sites respectively (Leone et al., 2008). Indeed, the same interaction has been recently studied by means of different experimental conditions, providing further elucidations about the structural details of the interaction interfaces (Lee et al., 2011). The binding mode occurring between Ship2-Sam and EphA2-Sam, and the type of stabilizing interactions, are very similar to those observed for the complex Odin-Sam1/EphA2-Sam. Indeed, the above mentioned NMR based-displacement

experiment (See Paragraph 2.3.3) shows that Ship2-Sam and Odin-Sam1 share a common binding site on the surface of the EphA2 receptor (**Figure 10**).

A comparison of our Odin-Sam1/EphA2-Sam model with the NMR structure of the tandem Sam domain of AIDA1b (PDB code: 2KIV, (Kurabi et al., 2009)) indicates analogies in between the two Sam-Sam associations. This tandem has a Mid-Loop/End-Helix topology in which AIDA1b-Sam1 and AIDA1b-Sam2 contributes the ML and EH binding surfaces respectively (Kurabi et al., 2009) (See Figure 6, Chapter 1). The high sequence homology in between the Sam domains of AIDA1b and Odin (**Figure 3**), suggests that in the tandem Odin Sam1-Sam2, the two domains may bind with an analogous ML/EH model in which Odin-Sam1 supplies the ML interface. Clearly, if in the full length Odin protein the ML surface of Odin-Sam1 is engaged in the interaction with Odin-Sam2, uncoupling of the two Sam domains from the tandem may take place to allow Sam1 binding to EphA2-Sam. Indeed, it has already been hypothesize opening of the AIDA1b-tandem to allow AIDA1b-Sam1 to interact with other proteins and explicate its functions (Kurabi et al., 2009).

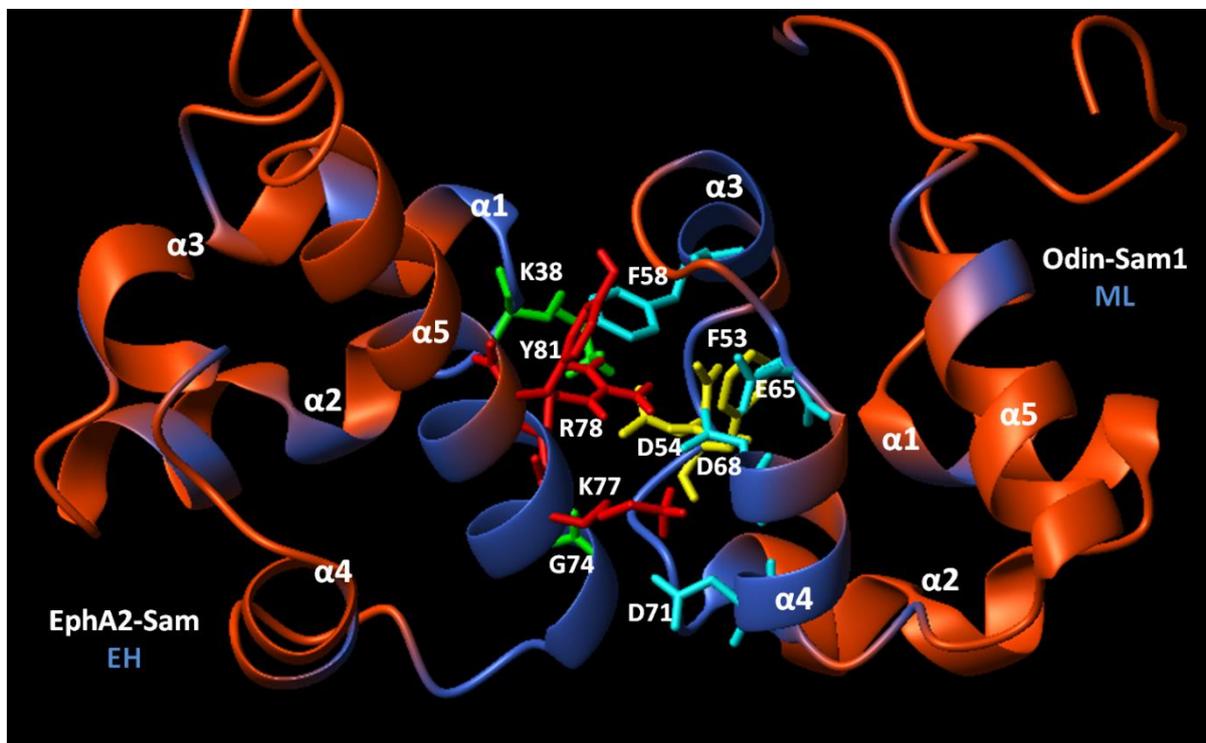


Figure 11. Haddock (de Vries et al., 2010) model (second ranked structure) of the EphA2-Sam/Odin-Sam1 complex, belonging to the most populated docking cluster. Protein binding interfaces, according to NMR chemical shift perturbation data, are colored blue, and side chains of a few residues providing interactions at the dimer interface are shown.

2.4. CONCLUSIONS

In this study, I have determined the Odin-Sam1 NMR solution structure and investigated the interaction between Odin-Sam1 and EphA2-Sam, demonstrating that the two domains bind with a dissociation constant in the low micromolar range, by forming a Mid-Loop/End-Helix hetero-dimer that looks highly stabilized by electrostatic interactions. In this model, the central part of Odin-Sam1 and the C-terminal helix, together with close loop regions of EphA2-Sam, supply the Mid-Loop and End-Helix binding sites respectively. Finally, I have implemented NMR-based displacement experiments and molecular docking studies to show that Ship2-Sam and Odin-Sam1 adopt very similar binding modes for the heterotypic Sam-Sam interactions with EphA2.

These data clarify molecular features related to an association involving the EphA2 receptor, which may be relevant in its endocytosis regulation during pathological processes. However more structural and biochemical studies are needed to shed light on the complex mechanisms regulating the network of interactions of Odin-Sam1 and EphA2-Sam.

It could be interesting to design, based on the structural information obtained here and in previous work, antagonists of Odin-Sam1/EphA2-Sam interaction which can modulate EphA2 receptor endocytosis and play a role as therapeutic agents.

3. STUDIES OF THE HETEROTYPIC SAM-SAM INTERACTION BETWEEN ODIN AND ARAP3

Sam (Sterile alpha motif) domains are small helical modules which mediate protein-protein interaction, and are made up of approximately 70 residues. In spite of a very similar fold, these domains are characterized by a high versatility regarding binding preferences and function (Kim and Bowie, 2003; Qiao and Bowie, 2005). Many Sam domain functions are mediated by formation of homo and heterotypic Sam-Sam associations (Meruelo and Bowie, 2009); in particular Sam domains can associate by forming dimers, oligomers and polymers (Meruelo and Bowie, 2009), by adopting head-to-head, tail-to-tail, and even head-to-tail topologies (Smalla et al., 1999; Thanos et al., 1999; Kim et al., 2001; Leone et al., 2008; Rajakulendran et al., 2008; Leone et al., 2009). This last binding mode is called Mid-Loop (ML)/End-Helix (EH) because residues from the central part of one Sam domain (Mid-Loop Interface) form intermolecular contacts with residues from the C-terminal helix and close loop regions of another Sam domain (End-Helix Interface).

Associations between the Sam domain from the PI3K effector protein Arap3 (Arap3-Sam) and the lipid phosphatase Ship2 (Ship2-Sam) (Leone et al., 2009), and between the first Sam domain of Odin (Odin-Sam1) and the Sam domain from the EphA2 receptor (EphA2-Sam) (Mercurio et al., 2012), were previously characterized, and shown that these complexes may adopt the head-to-tail topology of binding. This study focuses on analysis of the Odin-Sam1/Arap3-Sam interaction.

3.1. BACKGROUND

Arap3 (ArfGAP with Rho GAP domain, Ankyrin repeat and PH domain 3) is a phosphatidylinositol 3 kinase (PI3K) effector protein (Krugmann et al., 2002). It works as a GTPase activating protein for Arf6 and RhoA, and is involved in biological processes related to formation of lamellipodia, cell adhesion and spreading, regulation of actin cytoskeleton (I et al., 2004; Krugmann et al., 2006). In particular, Arap3 has been correlated to developmental angiogenesis (Gambardella et al., 2010) and scirrhous gastric carcinoma (Yagi et al., 2011).

Arap3 contains in its primary sequence a Sam domain (Kim and Bowie, 2003). In a yeast two-hybrid screen two Sam domains containing proteins were identified as possible Arap3 regulators: Ship2 (Src homology 2 domain-containing phosphoinositide-5-phosphatase 2) and ANKS1 (ANKyrin repeat and Sam domain containing 1) (Raaijmakers et al., 2007). The NMR (Nuclear Magnetic Resonance) solution structure and the Sam-Sam interaction between Arap3 and Ship2 have already been characterized (Leone et al., 2009; Raaijmakers et al., 2007). Arap3-Sam and Ship2-Sam bind to each other with a dissociation constant in the nanomolar range (Raaijmakers et al., 2007), and protein-protein association is mainly stabilized by salt bridges (Leone et al., 2009). Docking studies, together with NMR and mutagenesis data, show that Ship2-Sam and Arap3-Sam interact by means of a head-to-tail topology of binding, called Mid-Loop (ML)/End-Helix (EH) model (Raaijmakers et al., 2007; Ramachander and Bowie, 2004; Stafford et al., 2001). To date detailed analyses on the interaction between Arap3-Sam and ANKS1 protein Sam domains have not been reported. However, I have studied the heterotypic association between Odin-Sam1 and EphA2-Sam (See Chapter 2) (Mercurio et al., 2012), that represents an interaction probably significant for EphA2 receptor

endocytosis (Kim et al., 2010). Odin is a protein belonging to the ANKS family that owns two Sam domains in tandem (Sam1 and Sam2) (Emaduddin et al., 2008). Due to the rather high sequence homology between EphA2-Sam and Arap3-Sam (~58%), I have investigated if Arap3-Sam could interact with Odin-Sam1. To verify and thoroughly characterize the binding of the two proteins, NMR, ITC (Isothermal Titration Calorimetry), SPR (Surface Plasmon Resonance), molecular docking, and mutagenesis studies have been conducted.

3.2. MATERIALS AND METHODS

3.2.1. Protein expression

Sam domains were expressed in *Escherichia Coli* as recombinant proteins, using constructs provided with a tail of six His and a thrombin cleavage site at the N-terminus. Genes were cloned into the PET15B plasmids and transformed using BL21-Gold (DE3) competent cells (Stratagene). PET15B-constructs encoding human Arap3-Sam wild-type (residues 4-68, UniprotKB/TrEMBL code: Q8WWN8), Arap3-Sam triple (H37D, R77D, R80D) mutant, Odin-Sam1 (residues 691-770 of human Odin, UniprotKB/TrEMBL code: Q92625) and EphA2-Sam (residues 901-976 of human EphA2, Swiss-Prot/TrEMBL code: P29317) were purchased from Celtek Bioscience (Nashville, TN).

Unlabeled proteins expression was achieved by growing bacteria at 37°C in LB medium until $OD_{600}=0.6$. Protein over-expression was induced by IPTG (isopropyl β -D-thiogalactopyranoside) at a concentration of 1 mM, overnight at 25°C. Expression of $^{15}\text{N}/^{13}\text{C}$ double labeled and ^{15}N uniformly labeled proteins was carried out in M9 minimal medium supported with 2 g/L of ^{13}C -Glucose and/or 0.5 g/L of $^{15}\text{NH}_4\text{Cl}$. Purification of His-tag provided proteins was performed on a nickel column with an AKTA Purifier FPLC apparatus (Amersham).

3.2.2. Backbone resonance assignments of the Odin-Sam1/Arap3-Sam complex

NMR experiments were performed at 25 °C on a Varian Unity Inova 600 MHz spectrometer equipped with a cold probe.

In order to obtain resonance assignments for the backbone H, N and $C\alpha$ atoms of the Odin-Sam1/Arap3-Sam complex, HNCA and 3D ^{15}N resolved- ^1H , ^1H NOESY (100 ms mixing time) spectra were analyzed, acquired with samples containing $^{15}\text{N}/^{13}\text{C}$ double labeled Odin-Sam1 (620 μM) in presence of unlabeled Arap3-Sam (~3 mM), or doubly labeled Arap3-Sam (600 μM) in presence of unlabeled Odin-Sam1 (1.2 mM). NMR samples (600 μL volumes) consisted of phosphate buffer saline (PBS, 10 mM phosphates, 138 mM NaCl, 2.7 mM KCl) (Fisher) at pH=7.7 with 0.2% NaN_3 and 5% D_2O . Spectra were processed with the Varian software (Vnmrj version 1.1D) and analyzed with NEASY, as implemented in the program Cara (Bartels et al., 1995) (<http://www.nmr.ch/>).

3.2.3. Relaxation measurements

The evaluation of backbone ^{15}N longitudinal (R1) and transversal (R2) relaxation rates were performed by means of NMR experiments acquired at 25 °C on a Varian Unity Inova 600 MHz spectrometer provided with a cold probe. Measurements were performed with two samples consisting of ^{15}N - ^{13}C double labeled Odin-Sam1 (620 μM) with unlabeled Arap3-Sam (3 mM), and ^{15}N - ^{13}C double labeled Arap3-Sam (100 μM) with unlabeled Odin-Sam1 (~ 300 μM).

R1 and R2 relaxation data were collected as 1D spectra (4 K data points and 1- 4 K transients), recorded with five relaxation delays for R1 measurements (0.01, 0.1, 0.3, 0.6, 1.0 s), and seven relaxation delays for R2 measurements (0.01, 0.03, 0.05, 0.07, 0.09, 0.11, 0.15, 0.19 s).

The decrease in signal intensity as function of relaxation delay allowed to estimate average R1 and R2 values; then average R1/R2 ratios were used as input for the software timest (A. G. Palmer III, Columbia University) (Kay et al., 1989) in order to calculate the rotational correlation time (τ_c).

3.2.4. Odin-Sam1/Arap3-Sam binding studies

To study the protein-protein interaction chemical shift perturbation studies (Pellecchia, 2005) were performed, recording 2D [^1H , ^{15}N] HSQC experiments. The Odin-Sam1 binding interface for Arap3-Sam was evaluated by analysis of 2D [^1H , ^{15}N] HSQC spectra of ^{15}N labeled Odin-Sam1 (80 μM) in the unbound-form and after addition of unlabeled Arap3-Sam (120 μM , 370 μM , 620 μM). To map the binding site of Arap3-Sam for Odin-Sam1 2D [^1H , ^{15}N] HSQC spectra of a ^{15}N labeled Arap3-Sam sample (90 μM), in absence and presence of unlabeled Odin-Sam1 (100 μM , 300 μM and 800 μM), were acquired.

Similar NMR chemical shift perturbation experiments were conducted to assess the binding of Arap3-Sam (H37D, R77D, R80D) mutant to Odin-Sam1, by using ^{15}N labeled Arap3-Sam triple mutant (200 μM) and unlabeled Odin-Sam1 (200 and 300 μM).

An NMR displacement experiment was also carried out, in which 2D [^1H , ^{15}N] HSQC spectra were recorded for a ^{15}N labeled Arap3 protein sample (50 μM) in its unbound form, after addition of Odin-Sam1 (150 μM), and in presence of both Odin-Sam1 and EphA2-Sam (final protein ratio Arap3-Sam/Odin-Sam1/EphA2-Sam is 1: 3: ~40). Analyses of titration experiments and overlays of 2D spectra were obtained with the program Sparky (T. D. Goddard and D. G. Kneller, SPARKY 3, University of California, San Francisco).

3.2.5. SPR experiments

SPR experiments were carried out at 25 °C and a constant 20 $\mu\text{L}/\text{min}$ flow rate, using as running buffer a solution of Hepes 10 mM, pH=7.4, NaCl 150 mM, surfactant P20 0.05% v/v (90 μL injected for each experiment). Arap3-Sam was immobilized in 10 mM acetate buffer pH 5.0 (flow rate 5 $\mu\text{L}/\text{min}$, time injection 7 min) on a CM5 Biacore sensor chip, using EDC/NHS chemistry (Johnsson et al., 1991), while Odin-Sam1 was used as analyte at different concentrations in the range of 0.2-40 μM . Residual reactive groups were deactivated with 1 M ethanolamine hydrochloride, pH=8.5; the reference channel was prepared by activating with EDC/NHS and deactivating with ethanolamine. Immobilization level for Arap3-Sam was 1840 RU. The BIA evaluation analysis package (version 4.1, GE Healthcare, Milano, Italy) was used to subtract the signal of the reference channel and to estimate K_D (Dissociation constant) values. RU_{max} (Resonance Units) values of each experiment versus Odin-Sam1 concentration were fit by non-linear regression analysis with GraphPad Prism, version 4.00 (GraphPad Software, San Diego, California) (Rich et al., 2001).

3.2.6. ITC experiments

ITC experiments were performed with an iTC200 calorimeter (Microcal/GEHealthcare, Milan, Italy). Arap3-Sam (250 μM) was titrated into a solution of Odin-Sam1 (10 μM). Before ITC measurements both proteins were widely

dialyzed in the same buffer (PBS, pH 7.7). Similar ITC studies were conducted with a solution of Arap3-Sam (H37D, R77D, R80D) triple mutant (250 μ M in PBS pH=7.7) and Odin-Sam1 (10 μ M in PBS pH=7.7).

Fitting of data to a single binding site model was carried out with the Origin software as supplied by GE HealthCare.

3.2.7. Docking studies

Docking studies were performed through the Haddock web server (de Vries et al., 2010). To generate 3D models of the Odin-Sam1/Arap3-Sam complex, the first conformer of Odin-Sam1 NMR structures (PDB code: 2LMR (Mercurio et al., 2012)) was used and first five conformers of Arap3-Sam NMR ensemble (PDB code: 2KG5 (Leone et al., 2009)). The C-terminal flexible tail of Arap3-Sam (corresponding to residues 90-100) was not included in the docking procedure. Ambiguous interaction restraints were generated from chemical shift perturbation data: for both Odin-Sam1 and Arap3-Sam I set as active and passive those amino acids with higher normalized chemical shift deviations, and large solvent exposure, or which could potentially supply important intermolecular contacts, as revealed by structural homologies with other Sam-Sam complexes (for example the AIDA-1b tandem Sam domain, PDB code: 2KIV (Kurabi et al., 2009)). Hence, Odin-Sam1 residues set as active are L50, N51, F53, D54, D55, F58, E65, E66, D68, D71, whereas passive residues are L49, G52, V56, H57, S61, N62, V63, M64, Q67, R70. For Arap3-Sam I set residues G73, H74, K76, R77 as active, and residues R75, L79 as passive. Regions from L50 to E66 of Odin-Sam1 and T72-R80, V36-L38 of Arap3-Sam were set as semi-flexible. Finally, the N and C-terminal tail of Odin-Sam1 and the N-terminal tail of Arap3-Sam were put as fully flexible during all the docking stages.

The solvated docking mode was implemented (van Dijk and Bonvin, 2006). During the first phase of the docking protocol (the rigid body energy minimization) 1000 structures were generated; in the second step the best 200 solutions were subjected to semi-flexible simulated annealing, then a final refinement in water was performed. During a first selection, the final 200 Haddock solutions were visually analyzed with MOLMOL (Koradi et al., 1996), and models appearing in divergence with experimental NMR data or presenting non canonical Sam-Sam orientations, were discarded. The 106 selected Haddock models were inspected with MOLMOL (Koradi et al., 1996), and compared with experimental structures of other heterotypic Sam-Sam complexes. At the end of this screening, 23 models were chosen as representative of the possible conformations of the Arap3-Sam/Odin-Sam1 complex. The selected ensemble was clustered using an RMSD (Root Mean Square Deviation) cutoff value of 2.6 Å by using MOLMOL (Koradi et al., 1996). This clustering procedure was performed by superimposing the models on the backbone atoms of the secondary structure elements of both Arap3-Sam (residues 29-34, 39-47, 61-66, 72-83) and Odin-Sam1 (residues 32-39, 42-49, 57-60, 66-72, 77-88), and indicated the presence of 4 clusters of structures (See Appendix 2).

3.3. RESULTS AND DISCUSSION

3.3.1. Odin-Sam1/Arap3-Sam interaction studies

Firstly, chemical shift perturbation studies with 2D [^1H , ^{15}N] HSQC experiments were conducted to verify the association between Odin-Sam1 and Arap3-Sam. NMR spectra of a ^{15}N uniformly labeled Odin-Sam1 sample in the unbound form and in presence of unlabeled Arap3-Sam were acquired and compared. Several changes occur in the HSQC spectra of Odin-Sam1 recorded in presence of increasing amounts of Arap3-Sam, clearly indicating that an interaction takes place between these Sam domains (**Figure 1A**). Once reached saturation conditions (i.e., no more changes could be detected in the HSQC spectrum of Odin-Sam1 after further addition of Arap3-Sam), in order to identify the interaction surface of Odin-Sam1 the equation $\Delta\delta = [(\Delta H_N)^2 + (0.17 * \Delta^{15}\text{N})^2]^{1/2}$ (Farmer et al., 1996) was applied to evaluate proton and nitrogen normalized chemical shift deviations (**Figure 1B**). Largest variations, $\Delta\delta$ values ≥ 0.2 ppm, can be revealed in the central portion of Odin-Sam1, surrounding helices $\alpha 3$, $\alpha 4$ and the C-terminal portion of $\alpha 2$; changes also occur in the C-terminal flexible tail, presumably because residues from this region come in closeness with the main binding interface (**Figure 1C**).

The graph of chemical shift deviations against residue number obtained for Odin-Sam1 in complex with Arap3-Sam (**Figure 1B**), is very similar to the one of the protein in complex with EphA2-Sam (See Figure 5, Chapter 2).

The binding surface of Arap3-Sam for Odin-Sam1 was also mapped, using NMR experiments performed with ^{15}N labeled Arap3-Sam and unlabeled Odin-Sam1 (**Figure 2A**). Largest chemical shift deviations ($\Delta\delta$ values >0.1 ppm) affect the $\alpha 5$ helix and the close $\alpha 1\alpha 2$ and $\alpha 4\alpha 5$ loop areas (**Figure 2B, C**). Interestingly, the binding between Arap3-Sam and Ship2-Sam take place by means of analogues regions (Leone et al., 2009).

NMR perturbation data reveal for the Odin-Sam1/Arap3-Sam complex an head-to-tail topology of binding, where the central region of Odin-Sam1 and the $\alpha 5$ helix of Arap3-Sam, together with the adjacent loop regions, mainly provide the binding interfaces (**Figures 1-2**).

To establish the binding stoichiometry of Odin-Sam1/Arap3-Sam complex, ^{15}N longitudinal (R1) and transversal (R2) nuclear spin relaxation rates measurements were carried out: the evaluation of the R2/R1 average values enabled us to determine the correlation time (τ_c) of both Sam domains bound to each other (Farrow et al., 1994).

The τ_c of Arap3-Sam and Odin-Sam1 in their bound state result 10 ± 1 ns and 9.9 ± 0.8 ns respectively, these values are higher with respect to those evaluated for the proteins in their unbound forms (i.e., 7.3 ± 0.7 ns for Odin-Sam1 (Mercurio et al., 2012) and 8.2 ± 0.4 ns for Arap3-Sam (Leone et al., 2009)), due to the increased dimension of the complex and the consequent slower tumbling. Moreover, the τ_c estimates, ~ 10 ns, are comparable to those evaluated for Odin-Sam1/EphA2-Sam (Mercurio et al., 2012) and Ship2-Sam/Arap3-Sam (Leone et al., 2009) complexes (i.e. 11 ns), as well as the AIDA1b Sam1-Sam2 tandem (9.1 ns (Kurabi et al., 2009)) and indicate a 1:1 binding stoichiometry for the Odin-Sam1/Arap3-Sam interaction.

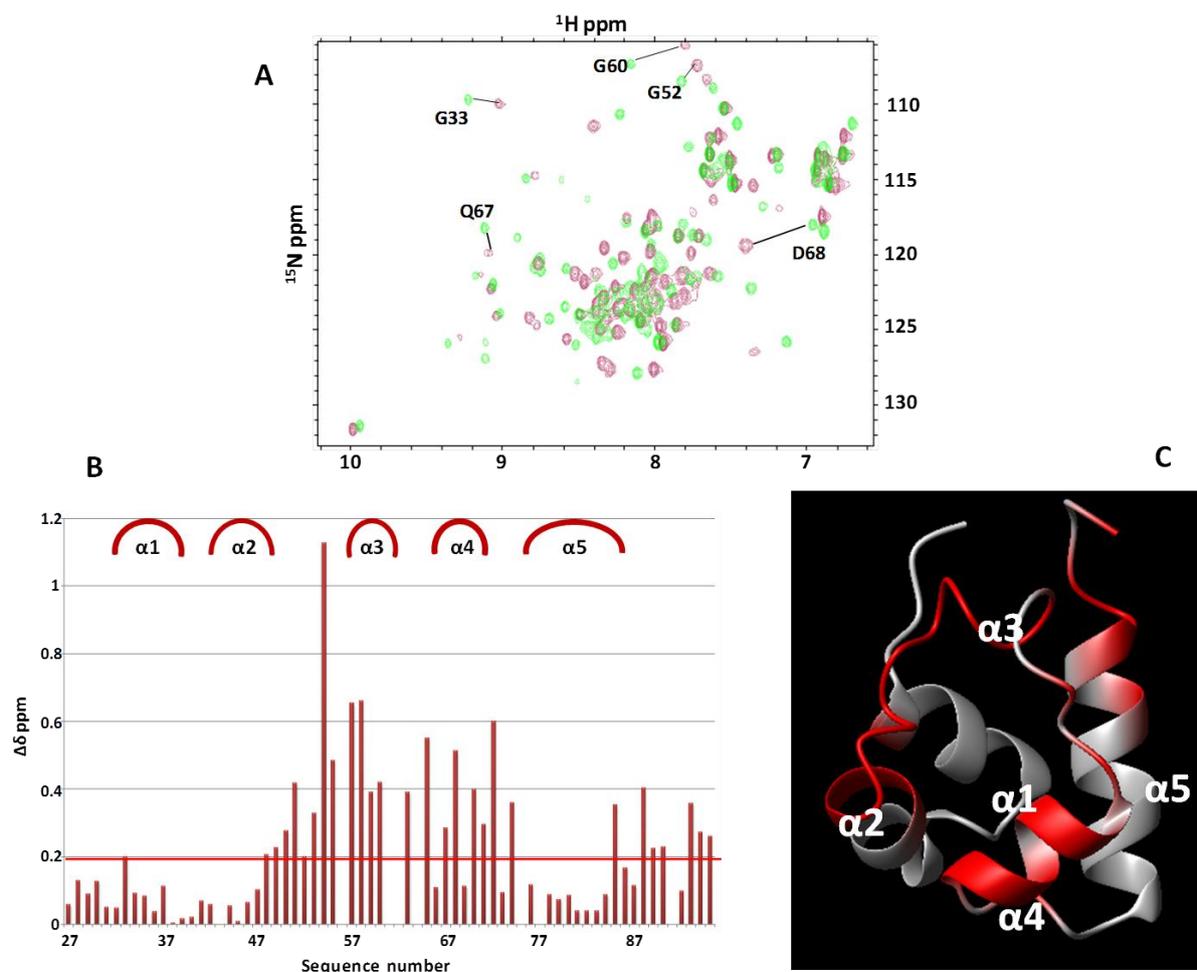


Figure 1. (A) Comparison of [^1H , ^{15}N] HSQC spectra of Odin-Sam1 (80 μM) in its unbound form (purple) and after addition of Arap3-Sam (370 μM) (green). (B) Histogram showing normalized chemical shift deviations as function of the residue number. Residues G33, L48, L49, L50, N51, G52, F53, D54, D55, H57, F58, L59, G60, V63, E65, Q67, D68, R70, D71, I72, I74, Q85, R88, S89, L90 V93, K94, A95 have normalized deviations ≥ 0.2 ppm. A $\Delta\delta$ equal to 0 has been assigned to residues Q43, V56, N62 whose peaks appear only in the spectrum of the complex, S61, M64 and S75 (unassigned), P77 and P91. (C) Residues with normalized chemical shifts deviation ≥ 0.2 ppm are represented in red on the 3D solution structure of Odin-Sam1 (PDB code: 2LMR (Mercurio et al., 2012)) in its ribbon representation.

To better investigate the Odin-Sam1/Arap3-Sam binding affinity, SPR and ITC experiments were performed. SPR binding assays (**Figure 3A**) were conducted by immobilizing Arap3-Sam on the chip surface and using Odin-Sam1 as analyte; a plot of RU_{max} values as function of Odin-Sam1 concentration, together with kinetic experiments, both implemented with a 1:1 binding model, gave a dissociation constant value (K_D) of 2.9 ± 0.4 μM . (**Figure 3B**).

ITC experiments (**Figure 3C**) resulted in rather good agreement with SPR studies and in fact, provided a $K_D = 0.37 \pm 0.08$ μM . ITC data indicated a single binding site model, as also shown by ^{15}N relaxation measurements. The calculated K_D values are similar to those determined in our previous SPR and ITC experiments for the Odin-Sam1/EphA2-Sam complex (5.5 ± 0.9 μM and 0.62 ± 0.04 μM through the two

techniques respectively) (Mercurio et al., 2012), suggesting comparable binding affinities of Arap3-Sam and EphA2-Sam for Odin-Sam1.

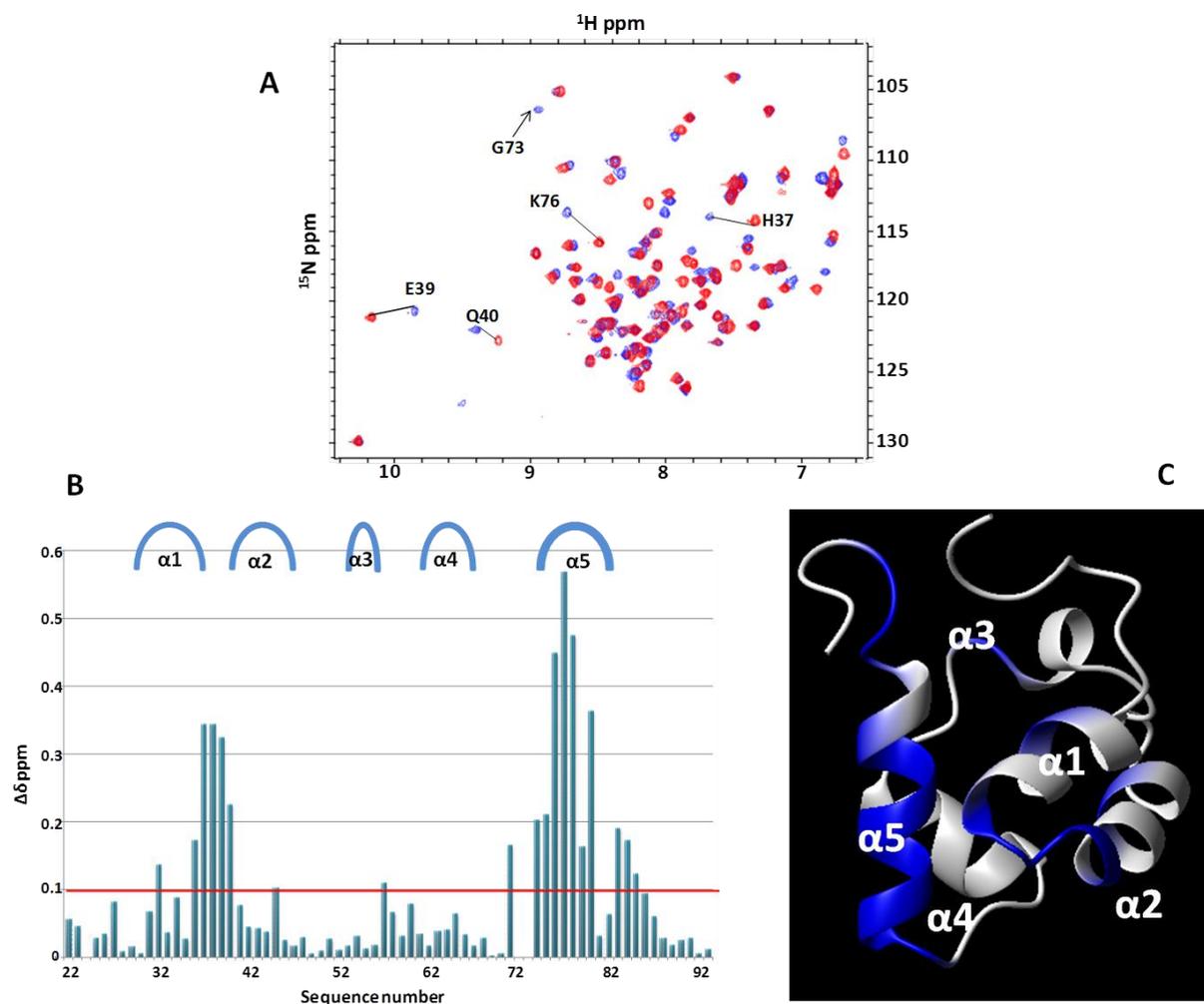


Figure 2. (A) Comparison of [^1H , ^{15}N] HSQC spectra of Arap3-Sam (90 μM) in its unbound form (red) and after addition of Odin-Sam1 (300 μM) (blue). (B) Histogram showing normalized chemical shift deviations as function of the residue number. W32, V36, H37, L38, E39, Q40, F45, R57, A71, H74, R75, K76, R77, I78, L79, R80, Q83, T84, G85 have deviations with values > 0.1 ppm. $\Delta\delta$ values equal to 0 have been attributed to residues T72, G73 (their peaks can be only seen in the spectrum of the complex), and P24. (C) Residues with normalized chemical shifts deviations > 0.1 ppm are represented in blue on the 3D solution structure of Arap3 (PDB code: 2KG5 (Leone et al., 2009)) in its ribbon representation.

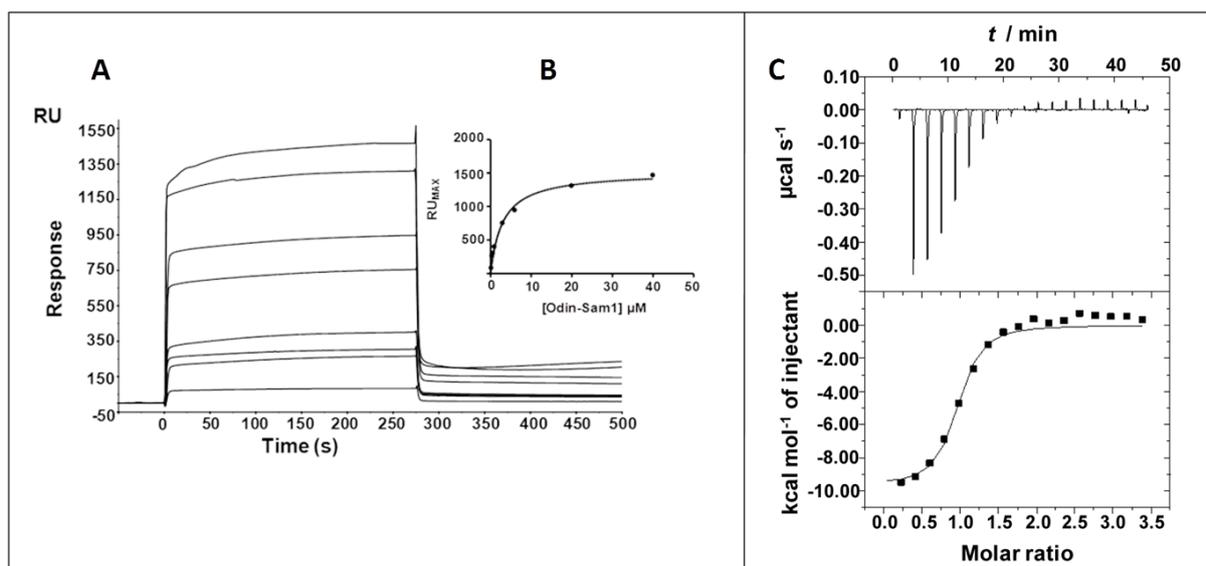


Figure 3. (A) SPR studies: overlay of sensorgrams representing the direct binding of Odin-Sam1 to immobilized Arap3-Sam (0.2-40 μM) (B) A plot of RU_{max} from each binding as function of Odin-Sam1 concentration. (C) ITC studies: calorimetric curve of the titration of Arap3-Sam (250 μM) with Odin-Sam1 (10 μM) is shown. The row and the integrated data are reported in the top and bottom sections respectively. For data fitting (bottom) a single-binding model was applied (Mercurio et al., 2012).

3.3.2. Molecular docking of the Odin-Sam1/Arap3-Sam complex

To assess the pattern of possible intermolecular interactions characterizing the Odin-Sam1/Arap3-Sam binding interface, I have performed docking studies with the Haddock web-server (de Vries et al., 2010) and built 3D models of the complex starting from chemical shift perturbation data and analysis of experimentally determined structures of other heterotypic Sam-Sam associations. The models suggest for the Odin-Sam1/Arap3-Sam complex a Mid-Loop/End-Helix topology, in which the Mid-Loop interface is provided by Odin-Sam1, whereas the End-Helix surface resides on Arap3-Sam (Figure 4).

Electrostatic interactions appear important for this association, since Arap3-Sam binding region is rich in positively charged residues, and, on the other hand, Odin-Sam1 surface contains many negatively charged amino acids (Figure 4).

To better assess this point, binding of the Arap3-Sam (H37D, R77D, R80D) triple mutant (Leone et al., 2009) to Odin-Sam1 was investigated. In this mutant the charge of three residues that are positioned in the presumed Arap3-Sam End-Helix binding site were reversed (Figure 5).

NMR and ITC experiments show that the triple mutant does not interact with Odin-Sam1 (Figure 6A, B), and let speculate that either the mutated residues are supplying crucial interactions at the Sam-Sam interface, and/or are causing conformational changes in the binding region that void association.

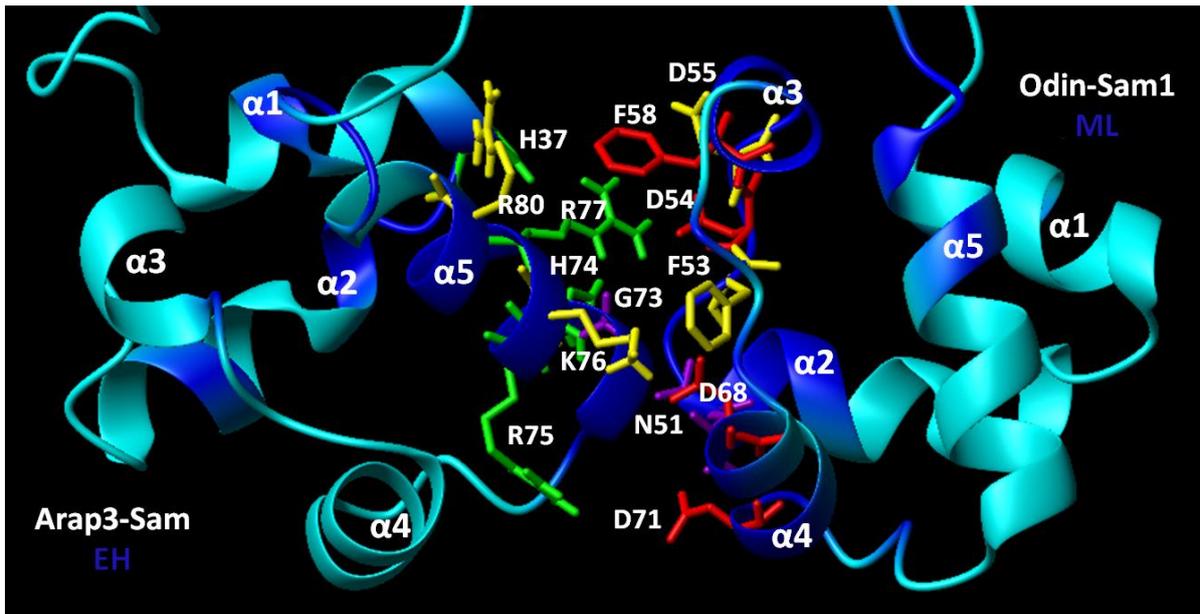


Figure 4. Haddock (de Vries et al., 2010) model (number 57) of the Odin-Sam1/EphA2-Sam complex, from the most populated docking cluster. Protein binding interfaces, according to NMR chemical shift perturbation data, are colored blue, and side chains of a few residues making contacts at the dimer interface are shown.

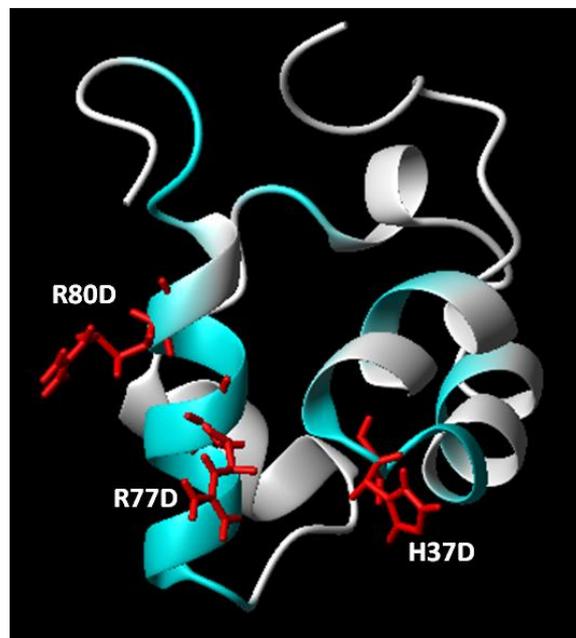


Figure 5. Ribbon representation of Arap3-Sam NMR solution structure (first conformer, PDB code: 2KG5 (Leone et al., 2009)), with residues corresponding to the binding interface for Odin-Sam1 highlighted in cyan. The side chain of amino acids replaced in Arap3-Sam triple mutant are shown in neon representation.

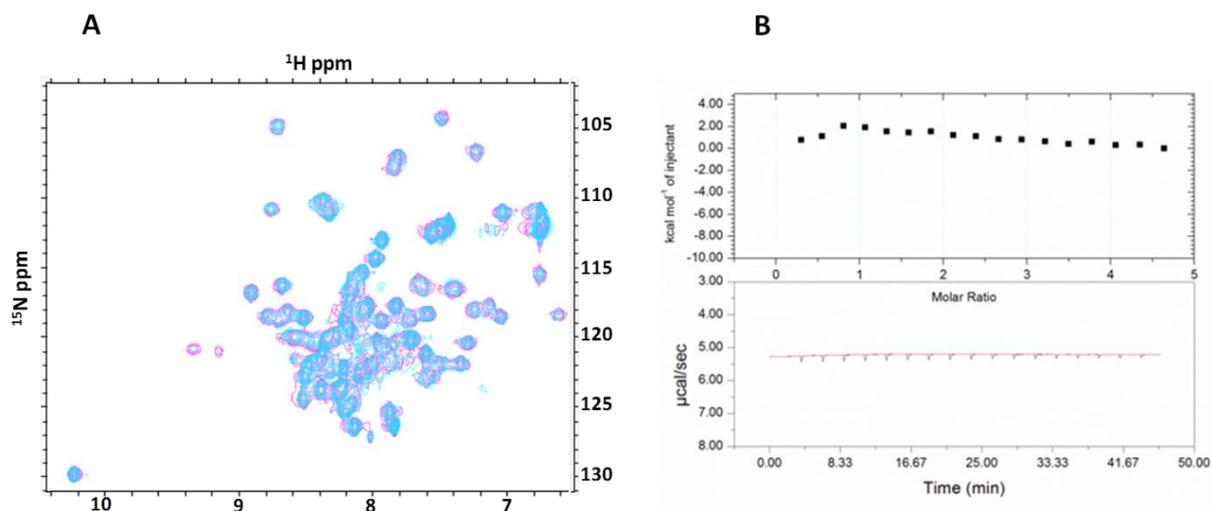


Figure 6. (A) Overlay of 2D [^1H , ^{15}N] HSQC spectra of ^{15}N labeled Arap3-Sam (H37D, R77D, R80D) triple mutant in absence (magenta) and in presence (cyan) of Odin-Sam1. **(B)** ITC data relative to Odin-Sam1 (10 μM) titration with Arap3-Sam triple mutant (250 μM) (Mercurio et al., 2013).

Mutagenesis data could be better interpreted by studying interactions pattern revealed in docking models.

Above all, Arg77 in most of the docking solutions (See Appendix 2) is engaged in electrostatic interactions with either Asp54 and/or Asp55; His37 and Arg80 seem less important. However, in a few models His37 makes a salt bridge with Asp55 and Arg80 may form a cation- π interaction with Phe58 or a salt bridge with Asp68 (See Appendix 2). The Arap3-Sam (H37D, R77D and R80D) mutant is also not capable of binding the Sam domain of the lipid phosphatase Ship2 (Leone et al., 2009), that instead can associate with wild-type Arap3-Sam with a ML/EH binding model in which Arap3-Sam furnishes the EH interface (Leone et al., 2009).

Another possible interaction on which to speculate in the Odin-Sam1/Arap3-Sam docking models is related to the residue Gly73, located on the EH surface of Arap3-Sam. The peak corresponding to the H_N of Gly73 can be only visualized in the [^1H , ^{15}N] HSQC spectrum of Arap3-Sam associated to Odin-Sam1 (**Figure 2A**), so indicating that this Gly, subsequent to complex formation, may become buried due to a possible engagement in an intermolecular contact. In fact, in a few docking solutions (See Appendix 2) Gly73 forms an intermolecular H-bond through its backbone H_N atom with the backbone C_O atom of Asn51 (**Figure 4**: Gly73 and Asn51 are colored magenta).

The interaction between the backbone amide proton of a Gly positioned at the base of the $\alpha 5$ helix on the EH interface, and the backbone carbonyl oxygen of another residue located at the C-terminal end of the $\alpha 2$ helix belonging to a ML surface, is present in a number of Sam-Sam complexes structures that have been experimentally determined (See PDB codes: 1PK1 (Kim et al., 2005), 2KIV (Kurabi et al., 2009), 3BS5 (Rajakulendran et al., 2008), 3SEI (Stafford et al., 2011), 3SEN (Stafford et al., 2011)). It has been speculated that a Gly residue in that position of an EH interface can facilitate the approach toward the ML binding region (Kurabi et al., 2009).

The Odin-Sam1/Arap3-Sam models reveal a resemblance to the docked structures of the Odin-Sam1/EphA2-Sam complex that are reported into our previous work (See Figure 11, Chapter 2), in which EphA2-Sam is providing the EH binding surface, and several intermolecular salt bridges occur at the Sam-Sam interface.

Moreover, a displacement assay performed with NMR techniques, indicated that Odin-Sam1 adopts the same ML binding mode to interact with both EphA2-Sam and Arap3-Sam (**Figure 7**), thus reflecting a similar behavior with respect to the Sam domain from the lipid phosphatase Ship2 (Ship2-Sam) (Leone et al., 2009). In fact, Ship2-Sam binds with its ML site both Arap3-Sam and EphA2-Sam (Leone et al., 2009; Leone et al., 2008).

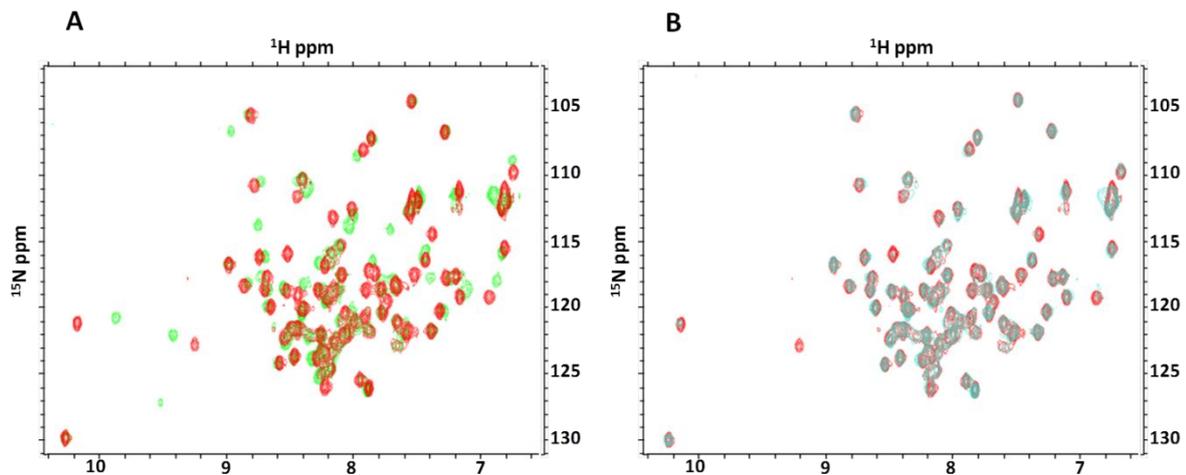


Figure 7 (A) Overlay of 2D [^1H , ^{15}N] HSQC spectra of ^{15}N labeled Arap3-Sam (50 μM) in the apo form (red) and bound to Odin-Sam1 (150 μM) (green). **(B)** The spectrum of unbound Arap3-Sam is re-established by addition of unlabeled EphA2-Sam (cyan).

The common structural features between heterotypic associations of Arap3-Sam and EphA2-Sam are probably a sign of the high sequence homology in between these two Sam domains (**Figure 8**), and point to a mutual protein interaction network.

Since EphA2 receptor endocytosis regulation (Kim et al., 2010; Zhuang et al., 2007) is related to Sam-Sam interactions between EphA2 and both Odin and Ship2, it cannot be excluded that Arap3-Sam is involved in this process by sequestering two essential regulators, but up to now the complex mechanism at the head of these heterotypic associations is not completely understood.

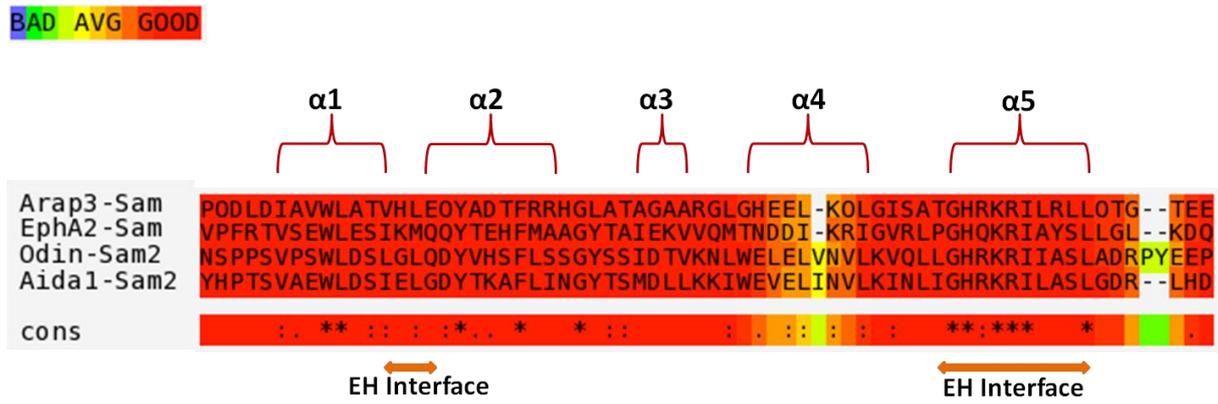


Figure 8. Sequence alignment of Arap3-Sam (Swiss-Prot/TrEMBL: Q8WWN8, residues 4-68) with EphA2-Sam (UniprotKB/TrEMBL code: P29317, residues 904-968), Odin-Sam2 (UniprotKB/TrEMBL code: Q92625, residues 770-837) and AIDA1b-Sam2 (UniprotKB/TrEMBL code: Q7Z6G8, residues 884-949). Secondary structure elements and regions corresponding to EH interface of Arap3-Sam are indicated with parenthesis and arrows respectively. Alignments have been generated with the T-COFFEE server (Notredame et al., 2000). Red residues correspond to highly reliable portions of the multiple alignment. The Cons line is a consensus, it indicates the average reliability value for every column.

3.4. CONCLUSIONS

In this project I have investigated the heterotypic Sam-Sam interaction between Odin and Arap3, revealing that the two proteins bind with a dissociation constant in the low micromolar range, a 1:1 stoichiometry, and implementing a structural model that closely resembles other ML/EH complexes, such as Ship2-Sam/Arap3-Sam (Leone et al., 2009) and Odin-Sam1/EphA2-Sam (Mercurio et al., 2012). In our model, the central part of Odin-Sam1 and the C-terminal helix together with adjacent loop regions of Arap3-Sam provide the ML and EH binding sites respectively.

Even if more experimental structural analyses need to be carried out, our preliminary results let hypothesize that interaction between Odin-Sam1 and Arap3-Sam may necessitate opening of the Odin Sam1-Sam2 tandem in the integral protein, as we have also supposed it is needed for the association between EphA2-Sam and Odin-Sam1 to occur (Mercurio et al., 2012).

Moreover, this report, in conjunction with earlier data, indicates for Arap3-Sam strict similarity to EphA2-Sam, in fact, both proteins bind with analogous structural topology either Ship2-Sam and Odin-Sam1 and share the same binding interfaces. Since Odin and Ship2 are both regulators of EphA2 receptor endocytosis (Kim et al., 2010; Zhuang et al., 2007), the exposed analogies might suggest a possible role of Arap3-Sam in this process. To date the functional consequences of the association between Arap3-Sam and Odin-Sam1 are unknown, thus our aim for the near future will be to generate new peptide/peptidomimetic molecules that could specifically antagonize Arap3-Sam heterotypic complexes, and be employed in cell-based assays to elucidate the functional implications of these interactions.

4. NMR STRUCTURAL CHARACTERIZATION OF THE SAM3 PEPTIDE

4.1. BACKGROUND

Odin is a member of the ANKS (Ankyrin repeat domain containing and Sam) protein family characterized by the presence of two Sam domains in tandem (Pandey et al., 2002).

Sam domains of ANKS proteins are potential inhibitors of tyrosine kinase EphA2 receptor degradation (Kim et al., 2010). It has been suggested a fine regulation of EphA2 receptor signaling by a balance between the activity of c-Cbl E3 ubiquitin ligase (a protein involved in Eph receptors degradation) and ANKS proteins (Kim et al., 2010).

The endocytosis and subsequent degradation of EphA2 and its connection with pro-cancer activities is under investigation; indeed enhanced receptor internalization and degradation are associated with decreased malignant cell behavior (Zhuang et al., 2007).

I have determined the Odin-Sam1 NMR (Nuclear Magnetic Resonance) solution structure and investigated the interaction between Odin-Sam1 and EphA2-Sam (Chapter 2) (Mercurio et al., 2012), demonstrating that the two domains bind with a dissociation constant in the low micromolar range, by forming a Mid-Loop/End-Helix hetero-dimer that looks highly stabilized by electrostatic interactions.

I have also investigated the heterotypic Sam-Sam interaction between Odin and the PI3K effector protein Arap3 (Chapter 3) (Mercurio et al., 2013); this study has revealed that the two proteins bind by implementing a structural model that closely resembles that occurring between Odin-Sam1 and EphA2-Sam (Chapter 2) (Mercurio et al., 2012).

On the basis of the structural knowledge acquired in the above mentioned studies, a peptide of 43 residues, named Sam3, which reproduces the sequence 715-758 of Odin-Sam1 (UniProtKB code Q92625) (**Figure 1A**) was designed. This peptide includes the central portion of Odin-Sam1 which is involved in the association with EphA2-Sam and Arap3-Sam, together with the $\alpha 5$ C-terminal helix (**Figure 1B, C**). This last region was added to induce the peptide to fold in a similar way with respect to the intact Odin-Sam1 domain.

Here, I report NMR studies on the conformational properties of Sam3 in PBS (Phosphate Buffer Saline) and H₂O/TFE (Trifluoroethanol) mixtures.

A **Sam3 sequence**
Ac-SKLLLN α GFDDVHFLGSNVMEEQDLRDIGISDPQHRRKLLQAAR-NH₂

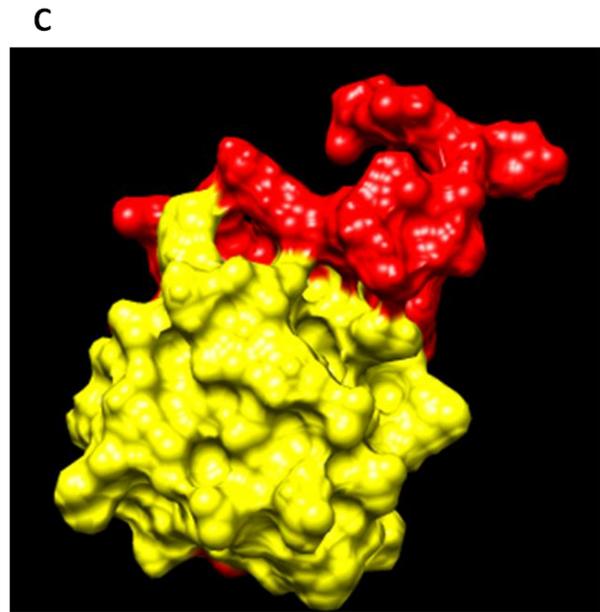
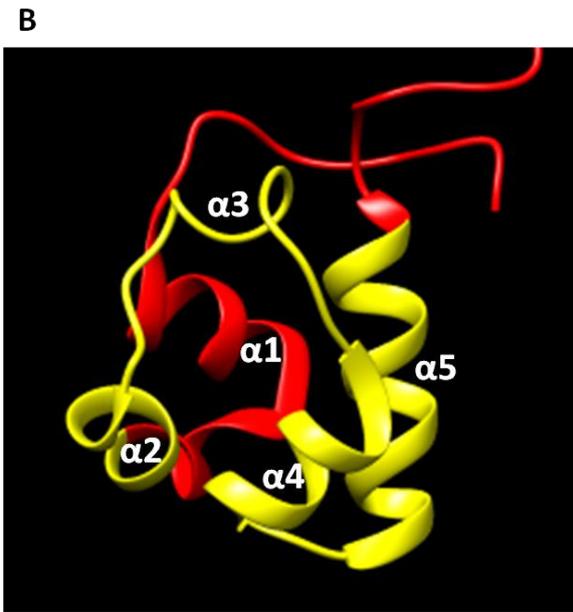


Figure 1. Sam3 primary sequence (A). Ribbon (B) and surface (C) representations of Odin-Sam1 with the region corresponding to the peptide highlighted in yellow.

4.2. MATERIALS AND METHODS

The acetylated-*N*-terminus and amidated-*C*-terminus peptide Sam3 was obtained in batch by Fmoc (Fluorenylmethyloxycarbonyl) standard chemistry protocol on Rink-amide 4-methylbenzhydrylamine (MBHA) resin. Peptide purification was achieved by RP-FPLC (Reversed Phase-High Performance Liquid Chromatography), and its integrity was confirmed by LC-MS (Liquid Chromatography-Mass Spectrometry) (Finnigan Surveyor, Thermo Electron Corporation).

NMR experiments were performed at 298 K on a Varian Unity Inova 600 MHz spectrometer equipped with a cold probe.

A first NMR sample was made up of Sam3 dissolved at a concentration of 730 μM in 550 μL of phosphate buffer saline (PBS, 10 mM phosphates, 138 mM NaCl, 2.7 mM KCl) (Fisher) at pH=7.7, plus 50 μL D₂O (98% D, Armar Scientific, Switzerland). For this sample, the following two dimensional (2D) NMR experiments were acquired: 2D [¹H, ¹H] TOCSY (Total Correlation Spectroscopy) (Griesinger et al, 1988), and 2D [¹H, ¹H] NOESY (Nuclear Overhauser Effect Spectroscopy) (Kumar et al., 1980).

2D proton NMR spectra were also recorded in TFE/H₂O (70/30 v/v). The sample was prepared by dissolving the peptide (630 μM concentration) in 420 μL TFE (2,2,2-Trifluoroethanol-D₃- 99.5% isotopic purity, Sigma-Aldrich) and 180 μL H₂O. For this sample, 2D [¹H, ¹H] TOCSY (Griesinger et al, 1988) (**Figure 2A**), 2D [¹H, ¹H] NOESY (Kumar et al., 1980) (**Figure 2B**), and 2D [¹H, ¹H] DQFCOSY (Double Quantum-Filtered Correlated Spectroscopy) (Piantini et al., 1982) were recorded. The process of proton resonance assignments was performed by following the Wüthrich protocol (Wüthrich,1986).

TSP (Trimethylsilyl-3-propionic acid sodium salt-D₄, 99% D, Armar Scientific, Switzerland) was used as internal standard for chemical shifts referencing (0.0 ppm).

2D experiments were generally acquired with 16-64 scans, 128-256 FIDs in the ω_1 dimension, 1024-2048 data points in the ω_2 dimension.

TOCSY spectra were acquired with 70 ms mixing time, NOESY experiments with 200 and 300 ms mixing times.

Water suppression was achieved with the DPFGE (Double Pulsed Field Gradient Selective Echo) sequence (Dalvit, 1998). Spectra were processed with VNMRJ 1.1D (Varian, Italy) and analyzed with NEASY (Bartels et al 1995) contained in the CARA software (<http://www.nmr.ch/>).

In order to evaluate temperature coefficients 2D [¹H, ¹H] TOCSY experiments were recorded at different temperatures: 298, 301, 304, 307, 310 K. A linear regression analysis of chemical shifts *versus* temperature was conducted (See Table 2, Appendix 3) with the program GraphPad Prism version 5.04 (GraphPad Software Inc., San Diego, USA).

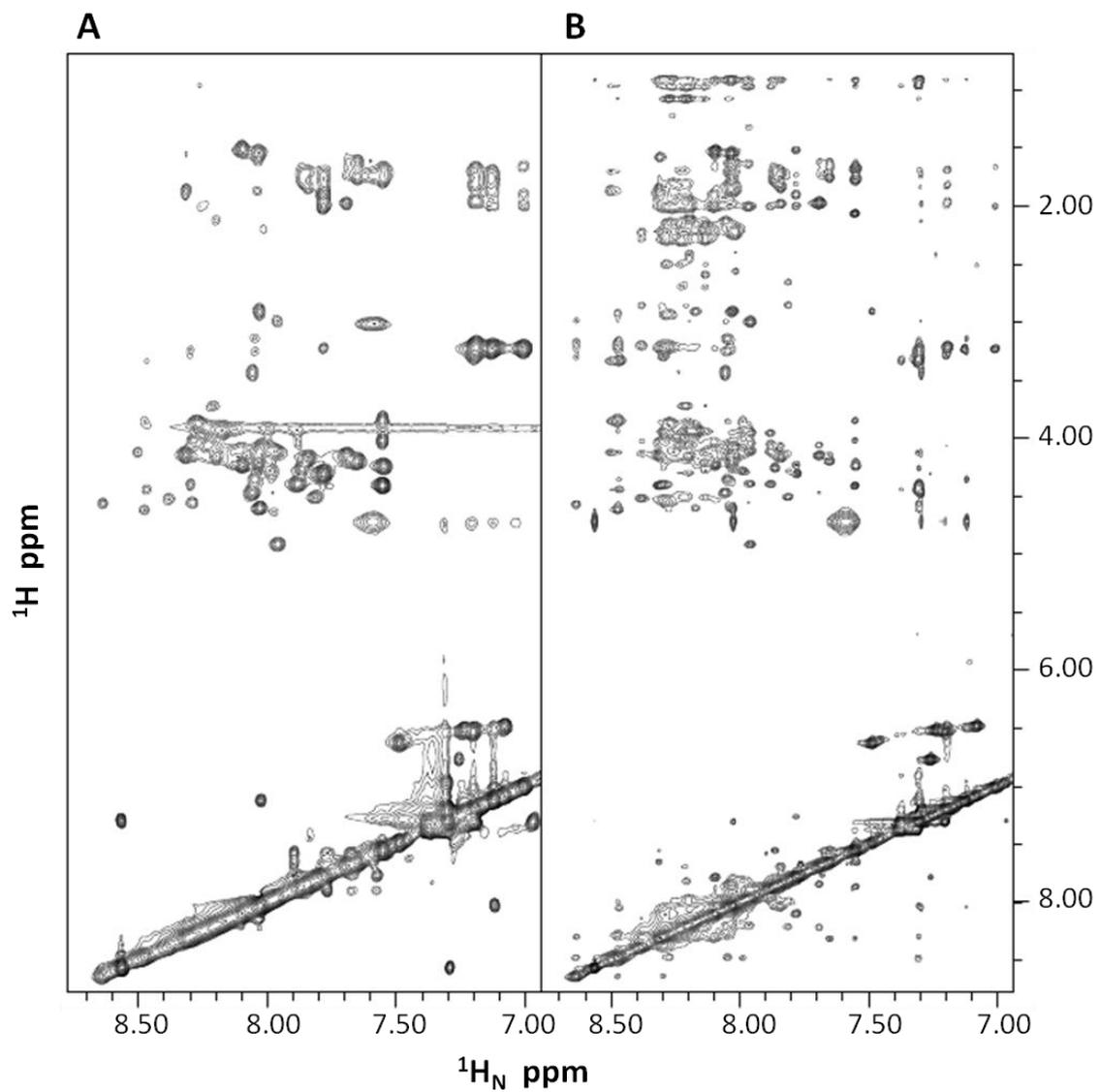


Figure 2. 2D [^1H , ^1H] TOCSY (A) and 2D [^1H , ^1H] NOESY (B) spectra of Sam3 recorded in $\text{H}_2\text{O}/\text{TFE}$, expansions containing signals from H_N and aromatic protons are showed.

4.3. RESULTS AND DISCUSSION

In my previous studies I have analyzed the molecular features at the basis of Odin-Sam1/EphA2-Sam and Odin-Sam1/Arap3-Sam complexes formation. The achieved results have been used to design a series of peptides reproducing Sam-Sam binding interfaces, with the purpose to evaluate their potential properties as antagonists of heterotypic Odin-Sam1 interactions.

The binding between two proteins is often characterized by a large interface, lacking a significant single binding pocket (Fletcher and Hamilton, 2006; Lo Conte et al., 1999), as occurs in the Odin-Sam1/EphA2-Sam and Odin-Sam1/Arap3-Sam associations (Mercurio et al., 2012; Mercurio et al., 2013). Thus, peptides, which are generally flexible and rather large molecules, are natural candidates in a rational design of biotechnological relevant compounds which, by mimicking the interacting area, could have inhibitory functions (Nieddu and Pasa, 2007; Parthasarathi et al., 2008).

The peptide Sam3 encompasses the central portion of the Odin Sam domain, which interacts with both EphA2-Sam and Arap3-Sam, and the $\alpha 5$ C-terminal helix (**Figure 1**). I have implemented NMR spectroscopy to study the solution structure of the peptide under different experimental conditions.

I have first investigated the conformational preferences of Sam3 in aqueous buffer (i.e., PBS).

In this solvent system the peptide is rather disordered and does not contain canonical secondary structure elements. In fact, the NOESY spectrum, acquired in PBS, has only a few cross peaks and is lacking NOE pattern characteristic of secondary structure elements (**Figure 3A**).

Further NMR experiments were recorded in presence of 70% TFE, which is frequently used as a folding agent in peptide conformational studies (Cammers-Goodwin et al., 1996). The effect of TFE on peptide conformation is most likely related to its ability to aggregate around the solute and exclude water, favoring the formation of intra-molecular hydrogen bonds and secondary structures (Roccatano et al., 2002).

Due to the high quality of the 2D spectra recorded for Sam3 in TFE/H₂O (70:30), I was able to achieve complete proton resonance assignments (Wüthrich, 1986). In detail, I have identified spin systems in 2D [¹H, ¹H] TOCSY and DQFCOSY spectra then, I have sequentially correlated them by analysis of the 2D [¹H, ¹H] NOESY experiments. In particular, NOESY spectra revealed the presence of NOE contacts attributable to helical secondary structures (**Figure 3B**): H_{Ni}-H_{Ni+1} cross-peaks were identified between amide protons H_N of all sequential residues (the pattern is obviously broken by Pro32); H α_i -NH_{i+3}, and H α_i -H β_{i+3} NOEs were mainly found in the regions Phe8-Arg25 and Gln33-Ala42, while H α_i -NH_{i+4} contacts were principally present in the segments Asn18-Arg25 and Gln33-Arg43 (**Figure 4**). In major detail, the NOE pattern points to an helix conformation in large part of the molecule, interrupted in the segment extending from Ile27 to Gln33, due to the presence of Pro32 (the Pro is an helix-breaker (MacArthur and Thornton, 1991)) (**Figure 4**). The contact H α_i -H δ_{i+1} in between the H α proton of Asp31 and the H δ protons of Pro32 indicates that the Pro is in *trans* configuration (Wüthrich, 1986).

Proton chemical shifts for the peptide Sam3 in TFE/H₂O (70/30) are reported in the Appendix 3.

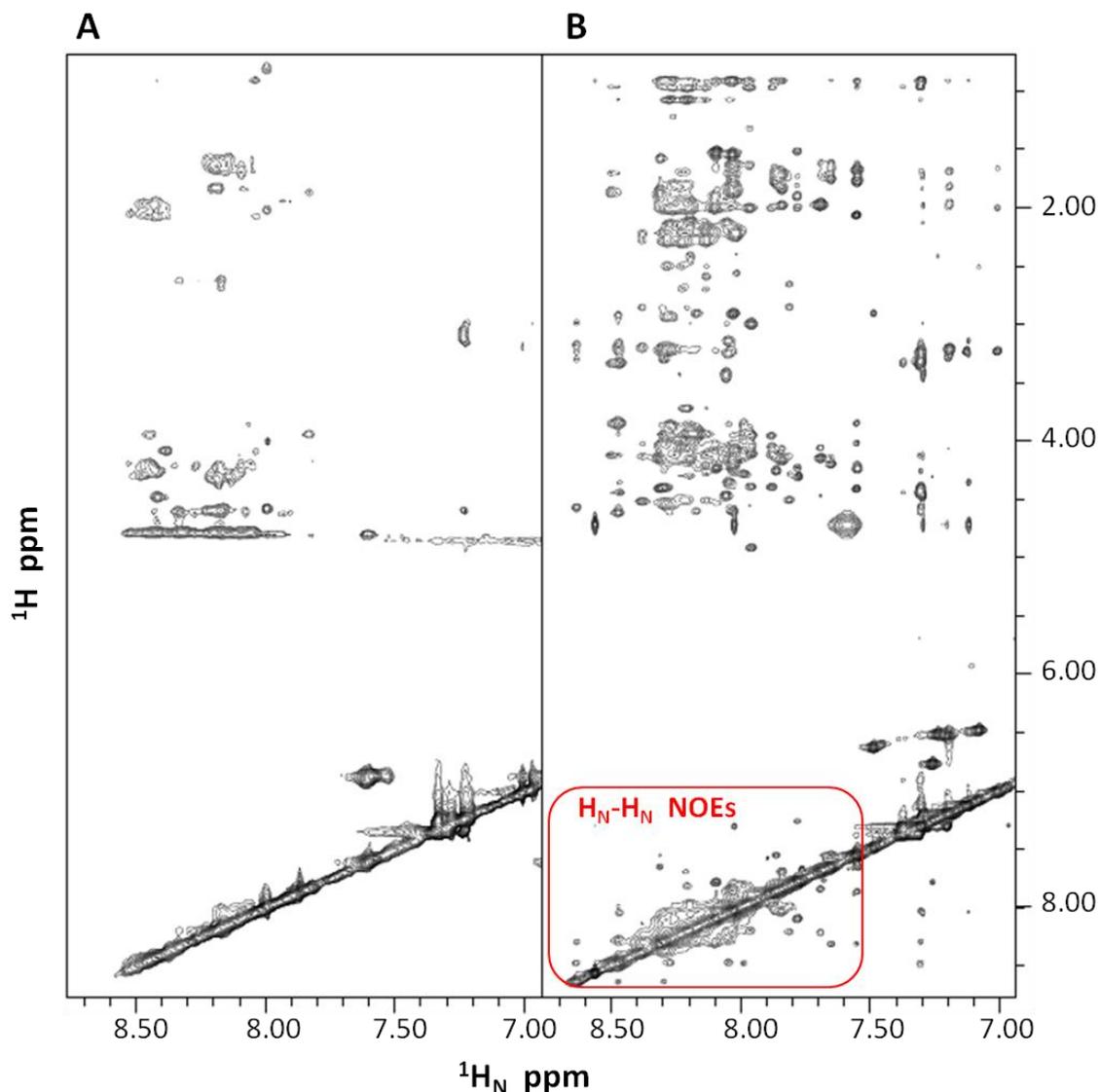


Figure 3. Comparison between Sam3 NOESY spectra recorded in PBS (**A**) and in TFE/H₂O (**B**). This last spectrum contains H_{Ni}-H_{Ni+1} NOE cross-peaks (inside the red rectangle), which are characteristic of helical secondary structures.

The Sam3 helical conformation, induced by TFE, is further supported by the α CH proton chemical shift deviations from random coil values (**Figure 5**).

NMR chemical shifts are highly sensitive probes for the local secondary structures of peptides and proteins (Wishart et al., 1991). In particular, those of α CH protons of all natural amino acids experience an upfield shift with respect to random coil values when in a helical configuration, and a downfield shift when in a β -strand extended configuration (Wishart et al., 1991).

As showed in the diagram of α CH chemical shift deviations versus residue number, almost all Sam3 residues present an upfield shift, thus further confirming helical secondary structures. The most evident exception (**Figure 5**) is present in the region Ile28-Pro32, where a helix interruption is likely to take place because of the Pro residue.

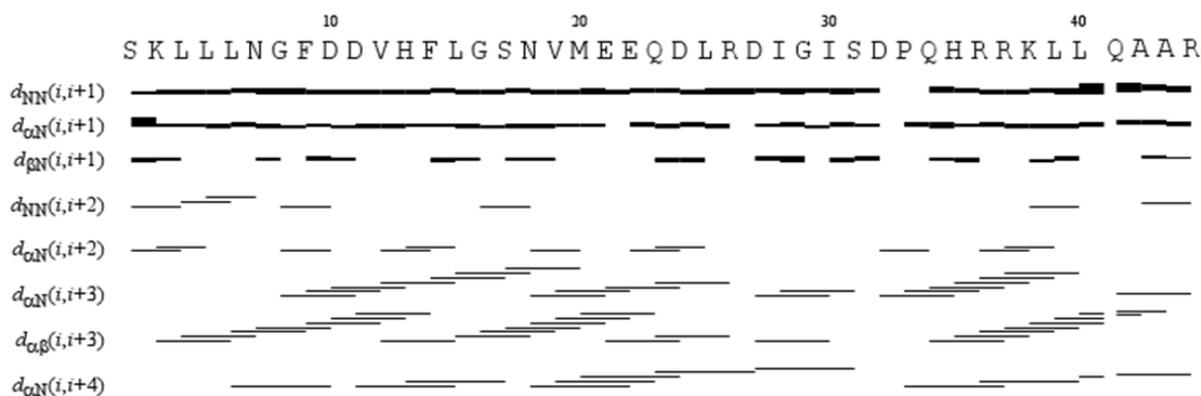


Figure 4. NOE pattern of Sam3 in TFE/H₂O.

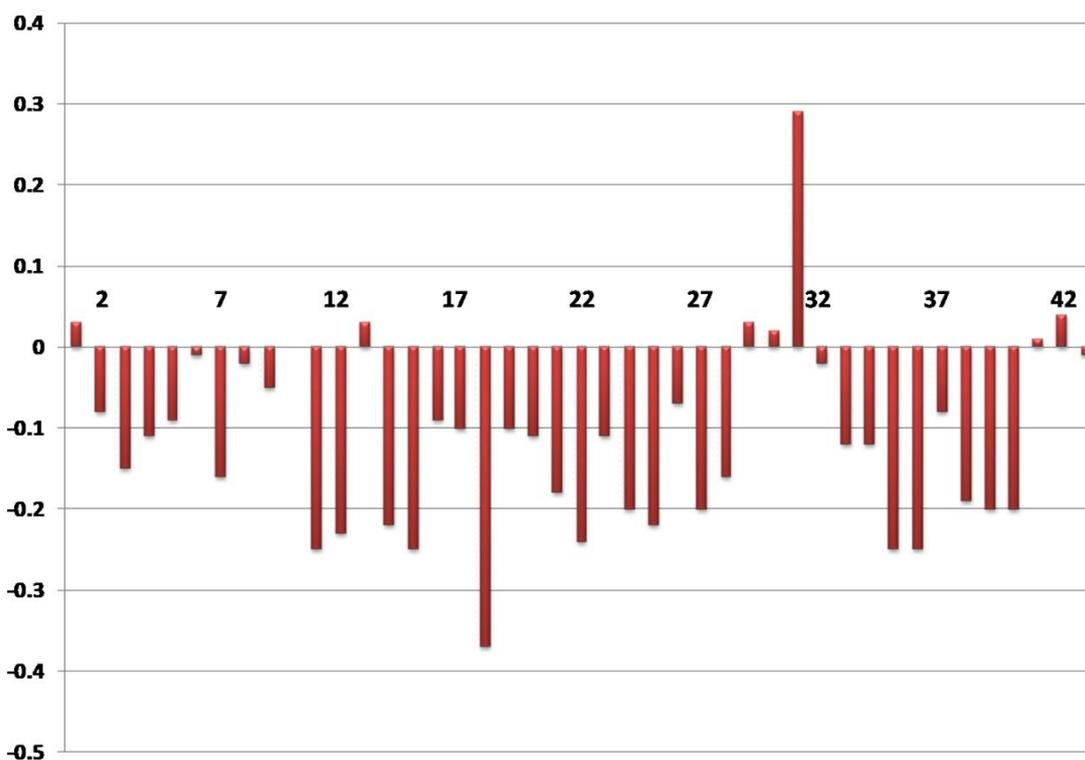


Figure 5. Chemical shift deviations from random coil values, evaluated in TFE/H₂O for Sam3 α CH protons (Wishart et al., 1991).

To assess the degree of solvent exposure of the Sam3 backbone amide protons in TFE, I have evaluated the temperature coefficients ($\Delta\delta/\Delta T$), by following chemical shifts variations as function of temperature (from 298 to 310 K) (Table 2, Appendix 3).

Chemical shifts of amide proton resonances have a temperature dependence; generally they shift upfield as the temperature increases (negative temperature

coefficients), and when the amide group is hydrogen-bonded because is involved in a structural motif, this effect occurs to a lesser extent (Baxter and Williamson, 1997). In proteins, generally temperature coefficient values more positive than -5 ppb/K indicate that the amide proton is involved in hydrogen bonding, whereas values ranging from -6 to -10 ppb/K suggest that the amide proton is solvent exposed (Baxter and Williamson, 1997).

$\Delta\delta/\Delta T$ values for most of the Sam3 amide protons at the C-terminal side (from Ser30 to Arg43) are included in the range from -1 to -5 ppb/K (Baxter and Williamson, 1997), thus indicating that H_N protons in this region have low solvent exposure and are likely hydrogen bonded. On the other hand, $\Delta\delta/\Delta T$ values more negative than -6 ppb/K are present in the stretch Phe8-Gly15, and may point towards a loss of secondary structure by increasing the temperature.

4.4. CONCLUSIONS

Within this work package I have investigated the conformational properties of a peptide, named Sam3, encompassing a large region of Odin-Sam1 including its binding sites for EphA2-Sam and Arap3-Sam.

The peptide structural analysis was carried out by means of NMR techniques.

Sam3 is disordered in PBS buffer, but after addition of TFE it shows a propensity to assume an helical structure. This result is supported by NMR methods which are predictive of peptides and proteins secondary structure: α CH chemical shift deviations from random coil values, NOE pattern.

Additional studies are ongoing to determine the Sam3 three-dimensional solution structure. Moreover, the ability of the Sam3 peptide to interact with Odin-Sam1 binding partners (EphA2-Sam and Arap3-Sam) and be implemented as a diagnostic or therapeutic tool, is under investigation.

The preliminary data acquired on the Sam3 peptide will be further used to design smaller and more constrained peptidomimetics. Our ultimate goal is to obtain molecule antagonists of Odin-Sam1/EphA2-Sam and Odin-Sam1/Arap3-Sam interactions that, by preserving some ordered secondary structure elements in aqueous buffer, could efficiently interfere with biological functions related to these Sam-Sam complexes.

5. CONCLUSIONS AND FUTURE PERSPECTIVES

In this PhD thesis, an interdisciplinary approach involving different experimental and computational techniques, such as NMR (Nuclear Magnetic Resonance), SPR (Surface Plasmon Resonance), ITC (Isothermal Titration Calorimetry), molecular docking, mutagenesis was undertaken, to thoroughly study the solution structure and the intermolecular interactions of a few Sam domains.

Sam domains are small protein modules made up of about 70-80 amino acids with a tertiary structure consisting of a five helix bundle; they are among the most abundant protein-protein interaction motives in the organisms (Qiao et al., 2005), and are characterized by high versatility as concerning their binding properties (Kim and Bowie, 2003).

My researches deal with the analysis of protein-protein interactions mediated by Sam domains, and the study of solution conformational features of peptides encompassing the corresponding interaction surfaces.

In the first work package (Chapter 2), the NMR solution structure of the N-terminal Sam domain of Odin (Odin-Sam1) was solved, and interaction studies with the Sam domain of the EphA2 receptor (EphA2-Sam) were performed.

The tyrosine kinase receptor EphA2 is able to regulate important biological processes related to cell motility (Egea and Klein, 2007; Himanen et al., 2007; Pasquale, 2005), and play a fundamental role in pathological processes such as cancer (Pasquale, 2008). In particular, the process of receptor endocytosis is widely studied for its possible involvement in tumoral processes, in which EphA2 is overexpressed (Pasquale, 2010; Surawska et al., 2004). Regulation of receptor endocytosis in a few cases takes place by means of proteins binding EphA2 through heterotypic Sam-Sam interactions (Kim et al., 2010; Zhuang et al., 2007). Among them, it was reported that the lipid phosphatase Ship2 (Src homology 2 domain-containing phosphoinositide-5-phosphatase 2) can inhibit the EphA2 receptor endocytosis, with a consequent increase in tumor cell invasiveness and metastatic potential (Zhuang et al., 2007). Previous NMR and ITC studies of the interaction between the Sam domain of Ship2 (Ship2-Sam) and EphA2-Sam were carried out (Leone et al., 2008), and showed that EphA2-Sam associates with Ship2-Sam with a high binding affinity, by adopting a topology common to several Sam-Sam complexes called "Mid-Loop/End-Helix" (Kurabi et al., 2009; Rajakulendran et al., 2008; Ramachander et al., 2004; Stafford et al., 2011). Furthermore, docking studies showed that this complex is mainly stabilized by electrostatic interactions between positively charged residues on the EphA2-Sam surface and negatively charged amino acids on the Ship2-Sam interface.

A recent work indicates that the EphA2 receptor endocytosis is also regulated by the ANKS (Ankyrin repeat domain containing and Sam) protein family (Kim et al., 2010). Among proteins belonging to this family, Odin is characterized by the presence of two Sam domains in tandem, and its first Sam domain presents high sequence homology (about 69%) with Ship2-Sam; so we decided to investigate the possible interaction between Odin-Sam1 and EphA2-Sam.

The first step of these studies was the characterization of the Odin-Sam1 solution structure, by means of 2D and 3D NMR spectra acquired with ^{15}N labeled and $^{15}\text{N}/^{13}\text{C}$ double labeled protein samples. Odin-Sam1 solution structure consists of five α helices, arranged in a canonical Sam domain fold.

Afterwards, binding studies with EphA2-Sam were carried out. SPR and ITC experiments revealed a low micromolar binding affinity between the two proteins. NMR chemical shift perturbation experiments were conducted to identify the interaction surfaces between Odin-Sam1 and EphA2-Sam. These data were used for docking studies, to build 3D models of the complex. These models exhibit some similarities with the EphA2-Sam/Ship2-Sam complex, and with other, previously described, Sam-Sam hetero-dimers (Kurabi et al., 2009; Rajakulendran et al., 2008; Ramachander et al., 2004; Stafford et al., 2011). In fact, even in this case the binding topology turns out to be the "Mid-Loop/End-Helix" one, in which Odin-Sam1 provides the Mid-Loop surface that interacts with the EphA2-Sam End-Helix interface; the complex is mainly stabilized by electrostatic interactions.

Models obtained by molecular docking were validated by mutagenesis experiments. In particular an EphA2-Sam triple mutant (K38A/R78A/Y81S) (Leone et al., 2008), in which the End-Helix region was perturbed with suitable amino acid replacements, presents lower ability to bind Odin-Sam1.

Odin-Sam1/EphA2-Sam models have been compared with NMR experimental structures deposited in the PDB (Protein Data Bank), in particular the tandem Sam domain of AIDA1b. AIDA1b, like Odin, is a protein belonging to the ANKS protein family (Kurabi et al., 2009). The AIDA1b Sam tandem presents a Mid-Loop/End-Helix topology in which the Sam1 and Sam2 domains are providing the Mid-Loop and End-Helix binding surfaces respectively (Kurabi et al., 2009). Because of the high sequence homology between AIDA1b and Odin (See Figure 3, Chapter 2), we assume that the two Odin Sam domains in tandem may adopt a similar ML/EH topology of interaction in which the ML interface is provided by Odin-Sam1. This scenario let speculate that opening of the Odin Sam1-Sam2 tandem domain may be required to allow association with EphA2-Sam.

The second work package focused on study of the interaction between Odin-Sam1 and Arap3-Sam. Arap3 (Arf GAP, Rho GAP, Ankyrin repeat and PH domains) is a protein involved in phosphoinositol-3-kinase (PI3K) signaling pathways and contains a Sam domain in its primary sequence (Krugmann et al., 2002). It was reported that among Sam domains containing proteins, both Ship2 and Odin can interact with Arap3-Sam (Raaijmakers et al., 2007). The Ship2-Sam/Arap3-Sam complex has already been characterized (Leone et al., 2009): the binding occurs with high affinity, through the "Mid-Loop/End-Helix" topology. Because of the high sequence homology between Arap3-Sam and EphA2-Sam (approximately 58%), the possible Odin-Sam1/Arap3-Sam complex formation was investigated. For this purpose, I have adopted techniques similar to those applied in the first work package. SPR and ITC experiments indicated a binding affinity between the two proteins in the low micromolar range. Subsequently, by NMR chemical shift perturbation experiments I have identified the putative binding interfaces in between the two Sam domains. Finally, these data were used to build 3D models of the Odin-Sam1/Arap3-Sam complex by computational docking calculations. The results indicate that these proteins associate by means of the "Mid-Loop/End-Helix" model, stabilized by electrostatic interactions, likewise to other complexes mentioned here.

The formulated hypotheses about the Arap3-Sam interaction surface against Odin-Sam1 were further validated with mutagenesis experiments. Indeed, a triple Arap3-Sam mutant (H37D, R77D, R80D) (Leone et al., 2009), characterized by amino acid substitutions in the "End-Helix" region, loses the ability to bind Odin-Sam1. These results let hypothesize that also the interaction between Odin-Sam1 and Arap3-Sam

may necessitate separation of the Odin Sam1-Sam2 tandem in the integral protein, as we have also supposed it is needed for the association between EphA2-Sam and Odin-Sam1.

Once analyzed the structural factors distinctive of Odin-Sam1/EphA2-Sam and Odin-Sam1/Arap3-Sam complexes, a few peptides of different sizes were designed, including the putative interaction regions between these Sam domains. Among them, the peptide Sam3 was analyzed in greater detail by means of NMR techniques. Sam3 is a peptide of 43 residues, reproducing the sequence 715-758 of Odin-Sam1 (UniProtKB code: Q92625), corresponding to the central portion of the domain which from my previous studies resulted involved in complexes formation with EphA2-Sam and Arap3-Sam, together with the α 5 C-terminal helix (See Figure 1, Chapter 4).

Sam3 was produced by solid phase synthesis; proton bi-dimensional NMR spectra were acquired to obtain resonance assignments of single residues.

NMR spectra recorded in phosphate buffer indicate that Sam3 is quite disordered and without canonical secondary structure elements. On the other hand, as revealed by further NMR experiments, Sam3 shows a propensity to adopt more ordered helical structures in presence of trifluoroethanol (TFE), which is frequently used as a folding agent in peptide conformational analysis (Cammers-Goodwin et al., 1996).

Ongoing studies are evaluating the possible interaction between Sam3 and the Odin-Sam1 binding partners, i.e. EphA2-Sam and Arap3-Sam.

Based on the knowledge acquired within this thesis, about structural features significant to heterotypic Sam-Sam complexes formation, in the near future libraries of biotechnological relevant molecules (peptides and peptido-mimetics), that could selectively interfere with Odin-Sam1/EphA2-Sam or Odin-Sam1/Arap3-Sam associations, will be designed. Particularly, antagonists of the Odin-Sam1/EphA2-Sam interaction are expected to modulate EphA2 receptor endocytosis and play a role as therapeutic agents in diseases such as cancer. These assumptions will be validated through appropriate cell-based assays.

With regard to the Odin-Sam1/Arap3-Sam complex, the designed molecules will be implemented in cellular assays to identify new biological functions related to Arap3-Sam.

APPENDIX 1: Docking solutions of the EphA2-Sam/Odin-Sam1 complex.

Cluster analysis of the docking solutions indicated the presence of 5 families of structures (See Chapter 2 for details about the clusterization protocol):

Cluster 1: structures n. 2, 5, 26, 55, 61, 71, 73, 75, 84, 91, 116, 118, 130

Cluster 2: n. 8, 23, 113

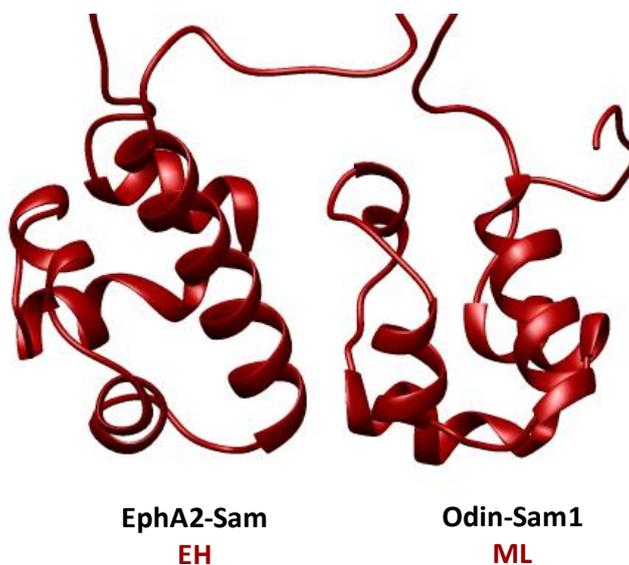
Cluster3: n. 16, 42, 109, 157

Cluster4: n. 20

Cluster5: n. 49

Structures are numbered according to the Haddock score (i.e., the lowest the better). Representative models from different conformational families are shown, together with lists of H-bonds and non-bonded contacts.

Structure n. 2. Haddock score=-129.34



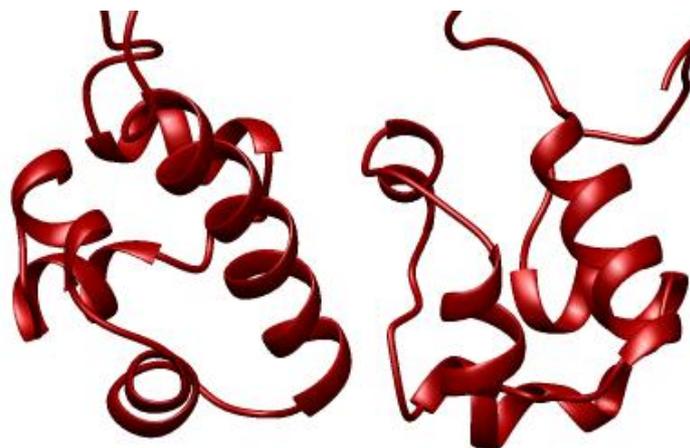
H-bond contacts

| EphA2-Sam | | | Odin-Sam1 | | | Distance (Å) |
|-----------------------|---------------------|-------------|-----------------------|---------------------|-------------|---------------------|
| Residue number | Residue type | atom | Residue number | Residue type | atom | |
| 36 | S | O | 61 | S | HG | 2.31 |
| 38 | K | HZ1 | 54 | D | OD1 | 1.65 |
| 38 | K | HZ3 | 55 | D | OD2 | 1.59 |
| 74 | G | HN | 50 | L | O | 2.45 |
| 77 | K | HZ1 | 68 | D | OD1 | 1.84 |
| 77 | K | HZ2 | 71 | D | OD1 | 1.90 |
| 77 | K | HZ3 | 67 | Q | O | 2.37 |
| 78 | R | HH12 | 55 | D | OD2 | 1.67 |
| 78 | R | HH22 | 55 | D | OD1 | 1.68 |

Non-bonded contacts

| EphA2-Sam | | | Odin-Sam1 | | | Distance (Å) |
|----------------|--------------|------|----------------|--------------|------|--------------|
| Residue number | Residue type | atom | Residue number | Residue type | atom | |
| 37 | I | CA | 58 | F | CE2 | 3.29 |
| 37 | I | CB | 58 | F | CE2 | 3.83 |
| 37 | I | CG2 | 58 | F | CE2 | 3.85 |
| 37 | I | C | 58 | F | CE2 | 3.66 |
| 38 | K | CG | 57 | H | CE1 | 3.51 |
| 74 | G | CA | 51 | N | CA | 3.18 |
| 74 | G | CA | 51 | N | CB | 3.85 |
| 74 | G | CA | 51 | N | CG | 3.35 |
| 74 | G | CA | 51 | N | C | 3.44 |
| 75 | H | CD2 | 52 | G | CA | 3.46 |
| 78 | R | CG | 58 | F | CZ | 3.61 |
| 78 | R | CD | 58 | F | CZ | 3.68 |
| 78 | R | CZ | 55 | D | CZ | 3.89 |
| 78 | R | CZ | 58 | F | CG | 3.84 |
| 78 | R | CZ | 58 | F | CD1 | 3.61 |
| 78 | R | CZ | 58 | F | CE1 | 3.63 |
| 78 | R | CZ | 58 | F | CZ | 3.86 |
| 81 | Y | CZ | 63 | V | CG1 | 3.80 |

Structure n. 71. Haddock score=-101.61



EphA2-Sam
EH

Odin-Sam1
ML

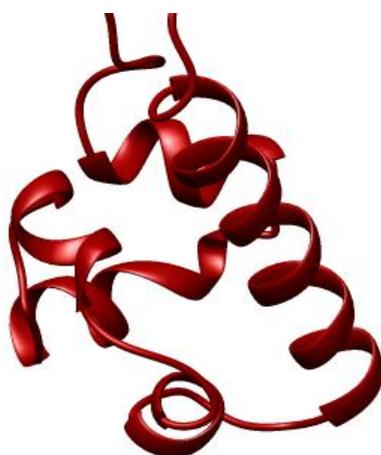
H-bond contacts

| EphA2-Sam | | | Odin-Sam1 | | | Distance (Å) |
|----------------|--------------|------|----------------|--------------|------|--------------|
| Residue number | Residue type | atom | Residue number | Residue type | atom | |
| 36 | S | O | 61 | S | HG | 2.09 |
| 38 | K | HZ1 | 55 | D | OD2 | 2.45 |
| 38 | K | HZ3 | 55 | D | OD1 | 1.92 |
| 38 | K | HZ3 | 55 | D | OD2 | 2.25 |
| 74 | G | HN | 51 | N | O | 2.30 |
| 77 | K | HZ1 | 68 | D | OD1 | 1.56 |
| 78 | R | HH12 | 55 | D | OD1 | 2.06 |
| 78 | R | HH21 | 52 | G | O | 1.78 |

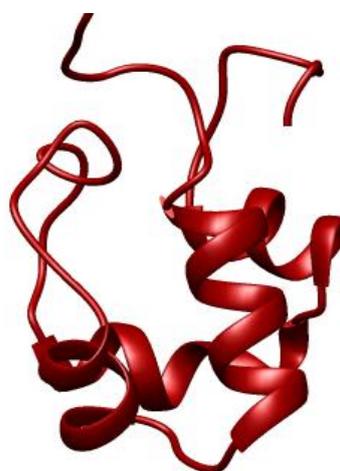
Non-bonded contacts

| EphA2-Sam | | | Odin-Sam1 | | | Distance (Å) |
|----------------|--------------|------|----------------|--------------|------|--------------|
| Residue number | Residue type | atom | Residue number | Residue type | atom | |
| 38 | K | CE | 57 | H | CB | 3.68 |
| 73 | P | CG | 51 | N | CA | 3.77 |
| 74 | G | CA | 51 | N | C | 3.81 |
| 74 | G | CA | 53 | F | CZ | 3.64 |
| 75 | H | CD2 | 52 | G | CA | 3.71 |
| 78 | R | CG | 58 | F | CE1 | 3.86 |
| 78 | R | CG | 58 | F | CZ | 3.52 |
| 78 | R | CZ | 58 | F | CG | 3.74 |
| 81 | Y | CE1 | 63 | V | CG1 | 3.90 |
| 81 | Y | CZ | 63 | V | CB | 3.89 |

Structure n. 8. Haddock score=-122.35



EphA2-Sam
EH



Odin-Sam1
ML

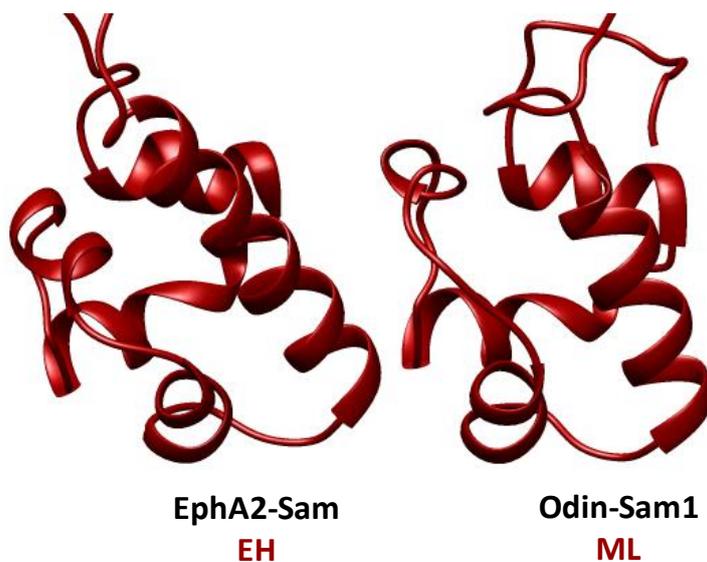
H-bonds

| EphA2-Sam | | | Odin-Sam1 | | | Distance (Å) |
|----------------|--------------|------|----------------|--------------|------|--------------|
| Residue number | Residue type | atom | Residue number | Residue type | atom | |
| 38 | K | HZ1 | 54 | D | OD1 | 1.94 |
| 38 | K | HZ2 | 55 | D | OD1 | 2.38 |
| 38 | K | HZ2 | 55 | D | OD2 | 1.58 |
| 77 | K | HZ1 | 68 | D | OD1 | 1.60 |
| 77 | K | HZ2 | 71 | D | OD1 | 1.69 |
| 77 | K | HZ3 | 67 | Q | O | 2.36 |
| 78 | R | HH12 | 55 | D | OD1 | 1.65 |
| 78 | R | HH21 | 52 | G | O | 1.70 |
| 78 | R | HH22 | 54 | D | OD2 | 2.47 |
| 78 | R | HH22 | 55 | D | OD1 | 2.29 |
| 81 | Y | HH | 61 | S | O | 1.80 |

Non-bonded contacts

| EphA2-Sam | | | Odin-Sam1 | | | Distance (Å) |
|----------------|--------------|------|----------------|--------------|------|--------------|
| Residue number | Residue type | atom | Residue number | Residue type | atom | |
| 38 | K | CE | 55 | D | CG | 3.89 |
| 74 | G | CA | 51 | N | CG | 3.73 |
| 74 | G | CA | 53 | F | CE2 | 3.86 |
| 75 | H | CD2 | 52 | G | CA | 3.82 |
| 77 | K | CD | 68 | D | CG | 3.84 |

Structure n. 16. Haddock score=-116.60



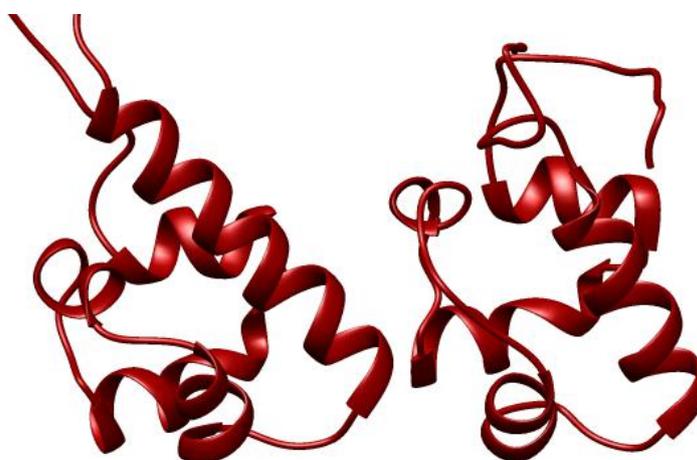
H-bonds

| EphA2-Sam | | | Odin-Sam1 | | | Distance (Å) |
|----------------|--------------|------|----------------|--------------|------|--------------|
| Residue number | Residue type | atom | Residue number | Residue type | atom | |
| 38 | K | HZ1 | 54 | D | OD1 | 1.62 |
| 38 | K | HZ1 | 54 | D | OD2 | 2.14 |
| 38 | K | HZ3 | 55 | D | OD2 | 1.56 |
| 74 | G | HN | 51 | N | OD1 | 2.30 |
| 77 | K | HZ3 | 68 | D | OD1 | 1.76 |
| 77 | K | HZ3 | 68 | D | OD2 | 2.49 |
| 78 | R | HH11 | 55 | D | OD1 | 1.78 |

Non-bonded contacts

| EphA2-Sam | | | Odin-Sam1 | | | Distance (Å) |
|----------------|--------------|------|----------------|--------------|------|--------------|
| Residue number | Residue type | atom | Residue number | Residue type | atom | |
| 73 | P | CG | 71 | D | CG | 3.87 |
| 74 | G | CA | 53 | F | CE2 | 3.87 |
| 74 | G | CA | 53 | F | CZ | 3.47 |
| 75 | H | CD2 | 51 | N | C | 3.44 |
| 75 | H | CD2 | 52 | G | CA | 3.48 |
| 78 | R | CA | 58 | F | CE1 | 3.84 |
| 78 | R | CZ | 58 | F | CB | 3.77 |
| 78 | R | CZ | 58 | F | CD1 | 3.57 |
| 81 | Y | CB | 58 | F | CE1 | 3.83 |
| 81 | Y | CB | 58 | F | CZ | 3.66 |
| 81 | Y | CD2 | 58 | F | CZ | 3.77 |
| 81 | Y | CD2 | 61 | S | CB | 3.88 |
| 81 | Y | CE2 | 61 | S | CB | 3.53 |

Structure n. 20. Haddock score=-114.57



EphA2-Sam

EH

Odin-Sam1

ML

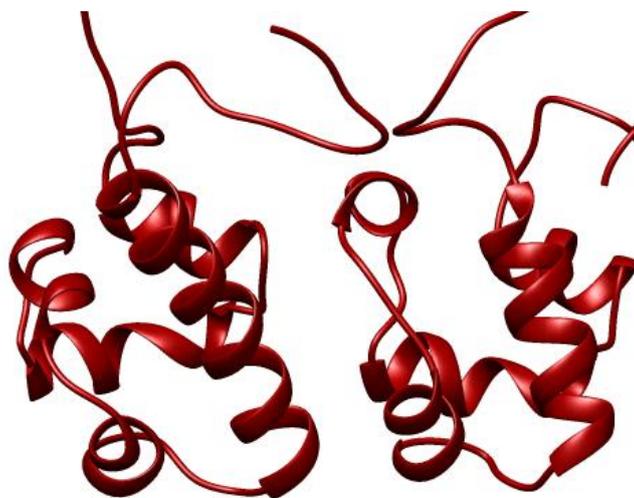
H-bonds

| EphA2-Sam | | | Odin-Sam1 | | | Distance (Å) |
|-----------------------|---------------------|-------------|-----------------------|---------------------|-------------|---------------------|
| Residue number | Residue type | atom | Residue number | Residue type | atom | |
| 38 | K | HZ1 | 54 | D | OD1 | 1.57 |
| 74 | G | HN | 51 | N | OD1 | 2.34 |
| 75 | H | HN | 51 | N | O | 2.04 |
| 77 | K | HZ1 | 61 | S | O | 2.16 |
| 77 | K | HZ2 | 62 | N | OD1 | 1.90 |
| 78 | R | HE | 55 | D | OD1 | 1.74 |
| 78 | R | HH22 | 54 | D | OD2 | 1.70 |
| 81 | Y | OH | 57 | H | HE2 | 2.39 |

Non-bonded contacts

| EphA2-Sam | | | Odin-Sam1 | | | Distance (Å) |
|-----------------------|---------------------|-------------|-----------------------|---------------------|-------------|---------------------|
| Residue number | Residue type | atom | Residue number | Residue type | atom | |
| 74 | G | CA | 52 | G | C | 3.68 |
| 74 | G | CA | 53 | F | CE2 | 3.90 |
| 74 | G | C | 52 | G | CA | 3.64 |
| 74 | G | C | 52 | G | C | 3.68 |
| 75 | H | CG | 52 | G | CA | 3.68 |
| 75 | H | CD2 | 52 | G | CA | 3.76 |
| 75 | H | CE1 | 50 | L | C | 3.74 |
| 75 | H | CE1 | 52 | G | CA | 3.76 |
| 77 | K | CB | 58 | F | CE1 | 3.65 |
| 77 | K | CB | 58 | F | CZ | 3.81 |
| 77 | K | CD | 62 | N | CG | 3.67 |
| 77 | K | C | 58 | F | CE2 | 3.83 |
| 77 | K | C | 58 | F | CZ | 3.44 |
| 78 | R | CA | 58 | F | CE2 | 3.61 |
| 81 | Y | CB | 58 | F | CZ | 3.32 |
| 81 | Y | CE2 | 57 | H | CD2 | 3.72 |
| 81 | Y | CZ | 57 | H | CD2 | 3.63 |

Structure n. 49. Haddock score=-106.39



EphA2-Sam
EH

Odin-Sam1
ML

H-bonds

| EphA2-Sam | | | Odin-Sam1 | | | Distance (Å) |
|-----------------------|---------------------|-------------|-----------------------|---------------------|-------------|---------------------|
| Residue number | Residue type | atom | Residue number | Residue type | atom | |
| 38 | K | HZ1 | 49 | L | O | 1.82 |
| 38 | K | HZ2 | 54 | D | OD2 | 1.61 |
| 74 | G | HN | 71 | D | OD1 | 2.29 |
| 77 | K | HZ1 | 68 | D | OD2 | 1.62 |
| 77 | K | HZ3 | 65 | E | OE2 | 1.58 |
| 78 | R | HH12 | 51 | N | OD1 | 1.73 |
| 78 | R | HH21 | 68 | D | OD1 | 1.63 |
| 81 | Y | HH | 61 | S | O | 2.13 |

Non-bonded contacts

| EphA2-Sam | | | Odin-Sam1 | | | Distance (Å) |
|-----------------------|---------------------|-------------|-----------------------|---------------------|-------------|---------------------|
| Residue number | Residue type | atom | Residue number | Residue type | atom | |
| 37 | I | CG2 | 58 | F | CE2 | 3.53 |
| 37 | I | CG2 | 58 | F | CZ | 3.64 |
| 38 | K | CD | 52 | G | CA | 3.32 |
| 78 | R | CZ | 53 | F | CE2 | 3.80 |
| 78 | R | CZ | 53 | F | CZ | 3.64 |

APPENDIX 2: Docking solutions of the Odin-Sam1/Arap3-Sam complex.

Docking solutions were subjected to a clusterization protocol (See Paragraph 3.2.7), from which 4 families of structures were obtained:

Cluster 1: structures n. 35, 45, 56, 57, 73, 89, 142, 166, 170, 175

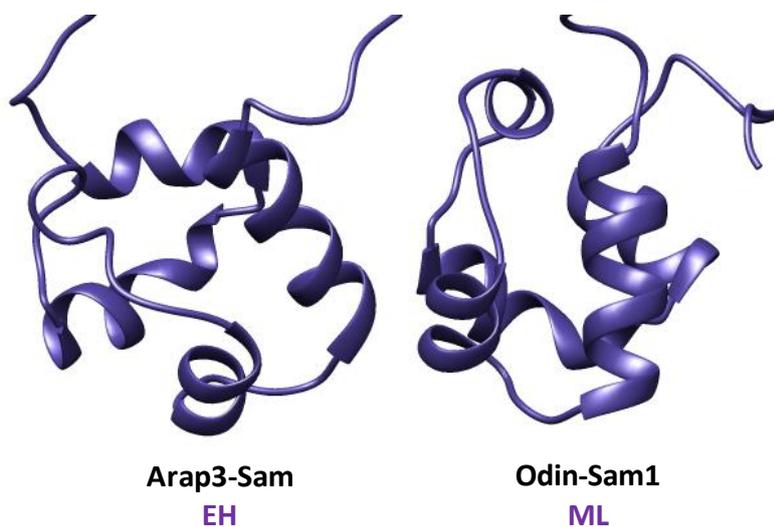
Cluster 2: n. 69, 71, 104, 105, 120, 139, 151, 196

Cluster 3: n. 94, 165, 185, 195

Cluster 4: n. 95

Structures are numbered according to the Haddock score (i.e., the lowest the better). The best models of each cluster, and corresponding tables of H-bond and non-bonded contacts, as determined by the Haddock web server (de Vries et al., 2010), are reported below.

Structure n. 35. Haddock score=-110.39



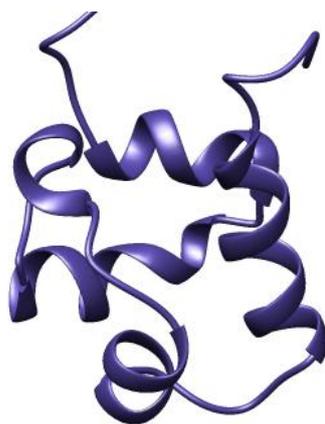
H-bond contacts

| EphA2-Sam | | | Odin-Sam1 | | | Distance (Å) |
|----------------|--------------|------|----------------|--------------|------|--------------|
| Residue number | Residue type | atom | Residue number | Residue type | atom | |
| 36 | S | O | 61 | S | HG | 2.31 |
| 38 | K | HZ1 | 54 | D | OD1 | 1.65 |
| 38 | K | HZ3 | 55 | D | OD2 | 1.59 |
| 74 | G | HN | 50 | L | O | 2.45 |
| 77 | K | HZ1 | 68 | D | OD1 | 1.84 |
| 77 | K | HZ2 | 71 | D | OD1 | 1.90 |
| 77 | K | HZ3 | 67 | Q | O | 2.37 |
| 78 | R | HH12 | 55 | D | OD2 | 1.67 |
| 78 | R | HH22 | 55 | D | OD1 | 1.68 |

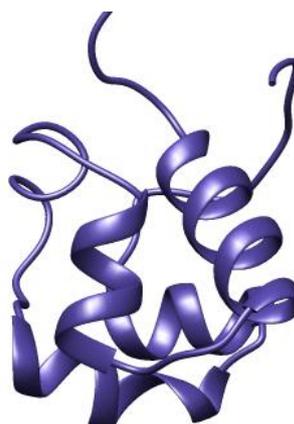
Non-bonded contacts

| Arap3-Sam | | | Odin-Sam1 | | | Distance (Å) |
|----------------|--------------|------|----------------|--------------|------|--------------|
| Residue number | Residue type | atom | Residue number | Residue type | atom | |
| 73 | G | CA | 51 | N | CA | 3.09 |
| 73 | G | CA | 51 | N | CG | 3.72 |
| 73 | G | CA | 51 | N | C | 3.43 |
| 73 | G | C | 51 | N | C | 3.88 |
| 74 | H | CE1 | 50 | L | CB | 3.81 |
| 74 | H | CE1 | 50 | L | CD2 | 3.81 |
| 76 | K | CE | 68 | D | CG | 3.76 |
| 77 | R | CZ | 52 | G | CA | 3.42 |
| 77 | R | CZ | 52 | G | C | 3.83 |

Structure n. 69. Haddock score=-100.42



Arap3-Sam
EH



Odin-Sam1
ML

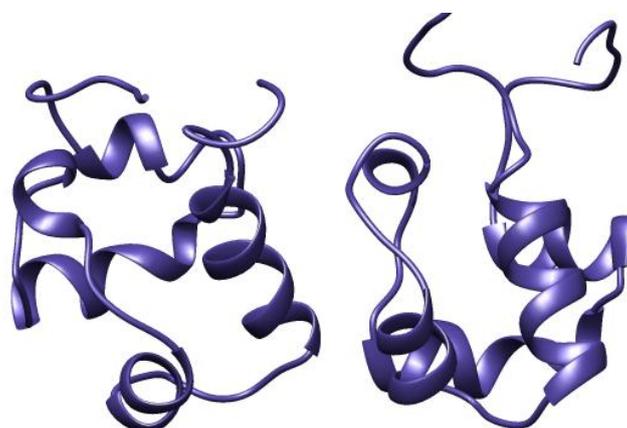
H-bond contacts

| Arap3-Sam | | | Odin-Sam1 | | | Distance (Å) |
|----------------|--------------|------|----------------|--------------|------|--------------|
| Residue number | Residue type | atom | Residue number | Residue type | atom | |
| 37 | H | HD1 | 55 | D | OD2 | 1.67 |
| 73 | G | HN | 50 | L | O | 2.35 |
| 74 | H | HE2 | 54 | D | OD2 | 2.38 |
| 76 | K | HZ1 | 71 | D | OD1 | 1.54 |
| 77 | R | HH12 | 55 | D | OD1 | 1.67 |
| 77 | R | HH22 | 52 | G | O | 1.90 |

Non-bonded contacts

| Arap3-Sam | | | Odin-Sam1 | | | Distance (Å) |
|----------------|--------------|------|----------------|--------------|------|--------------|
| Residue number | Residue type | atom | Residue number | Residue type | atom | |
| 73 | G | CA | 51 | N | CA | 3.49 |
| 73 | G | CA | 51 | N | C | 3.42 |
| 74 | H | CE1 | 52 | G | CA | 3.29 |
| 77 | R | CD | 58 | F | CE1 | 3.83 |
| 77 | R | CD | 58 | F | CZ | 3.77 |
| 77 | R | CZ | 58 | F | CD1 | 3.72 |
| 77 | R | CZ | 58 | F | CE1 | 3.62 |

Structure n. 94. Haddock score=-92.72



Arap3-Sam
EH

Odin-Sam1
ML

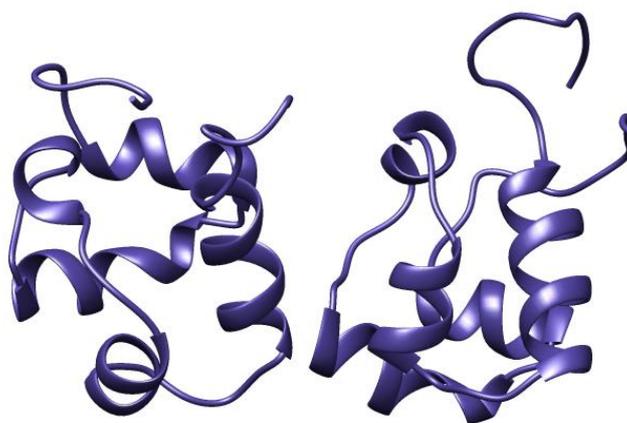
H-bond contacts

| Arap3-Sam | | | Odin-Sam1 | | | Distance (Å) |
|----------------|--------------|------|----------------|--------------|------|--------------|
| Residue number | Residue type | atom | Residue number | Residue type | atom | |
| 76 | K | HZ1 | 71 | D | OD1 | 1.67 |
| 76 | K | HZ2 | 68 | D | OD1 | 1.61 |
| 76 | K | HZ3 | 51 | N | OD1 | 1.83 |
| 77 | R | HH12 | 55 | D | OD2 | 1.80 |
| 77 | R | HH21 | 52 | G | O | 2.28 |
| 77 | R | HH22 | 55 | D | OD1 | 1.62 |

Non-bonded contacts

| Arap3-Sam | | | Odin-Sam1 | | | Distance (Å) |
|----------------|--------------|------|----------------|--------------|------|--------------|
| Residue number | Residue type | atom | Residue number | Residue type | atom | |
| 76 | K | CD | 53 | F | CE1 | 3.69 |
| 76 | K | CE | 71 | D | CG | 3.82 |
| 77 | R | CG | 58 | F | CE1 | 3.77 |
| 77 | R | CG | 58 | F | CZ | 3.58 |
| 77 | R | CZ | 55 | D | CG | 3.60 |
| 80 | R | CD | 63 | V | CB | 3.87 |

Structure n. 95. Haddock score=-92.71



Arap3-Sam
EH

Odin-Sam1
ML

H-bond contacts

| Arap3-Sam | | | Odin-Sam1 | | | Distance (Å) |
|----------------|--------------|------|----------------|--------------|------|--------------|
| Residue number | Residue type | atom | Residue number | Residue type | atom | |
| 74 | H | HE2 | 50 | L | O | 2.18 |
| 76 | K | HZ2 | 71 | D | OD2 | 1.57 |
| 77 | R | HH12 | 52 | G | O | 1.98 |
| 77 | R | HH21 | 51 | N | OD1 | 2.12 |
| 77 | R | HH22 | 52 | G | O | 2.15 |

Non-bonded contacts

| Arap3-Sam | | | Odin-Sam1 | | | Distance (Å) |
|----------------|--------------|------|----------------|--------------|------|--------------|
| Residue number | Residue type | atom | Residue number | Residue type | atom | |
| 37 | H | CE1 | 55 | D | CG | 3.90 |
| 74 | H | CD2 | 51 | N | C | 3.57 |

APPENDIX 3

Table 1. Chemical shifts list (± 0.01 ppm) of the peptide Sam3 evaluated at 298 K in TFE/H₂O (70:30).

| Residue | H _N | H α | H β | H γ | Others |
|-------------|----------------|------------|--------------|----------------------------|--|
| 1 S | 7.55 | 4.41 | 4.02 3.86 | | Acetyl 2.07 |
| 2 K | 8.231 | 4.14 | 1.91 1.86 | 1.59 1.51 | H δ 1.75 H ϵ 3.02 |
| 3 L | 7.65 | 4.20 | 1.76 1.64 | 1.69 | δ CH3 0.98, 0.92 |
| 4 L | 7.55 | 4.24 | 1.78 1.69 | 1.73 | δ CH3 0.97, 0.92 |
| 5 L | 7.87 | 4.27 | 1.78 1.71 | 1.79 | δ CH3 0.92 |
| 6 N | 8.17 | 4.61 | 2.93 2.91 | | H δ 7.49, 6.63 |
| 7 G | 8.17 | 3.95 | | | |
| 8 F | 8.30 | 4.41 | 3.31 3.25 | | H δ 7.25, 7.24 H ϵ 7.30 |
| 9 D | 8.63 | 4.57 | 3.18 3.00 | | |
| 10 D | 8.47 | 4.63 | 3.22 2.92 | | |
| 11 V | 8.27 | 3.87 | 2.16 | γ CH3 1.09, 0.92 | |
| 12 H | 8.04 | 4.36 | 3.25 3.15 | | H δ 7.12 H ϵ 8.03 |
| 13 F | 8.46 | 4.44 | 3.36 3.32 | | H δ 7.31 H ϵ 7.37 |
| 14 L | 8.50 | 4.13 | 1.89 1.71 | 1.87 | δ CH3 0.98 |
| 15 G | 8.47 | 3.86 | | | |
| 16 S | 7.98 | 4.29 | 4.02 | | |
| 17 N | 7.81 | 4.51 | 2.86 2.66 | | H δ 6.91, 5.70 |
| 18 V | 8.20 | 3.73 | 2.23 | γ CH3 1.09, 0.99 | |
| 19 M | 8.13 | 4.22 | 2.22 | 2.71 2.60 | |
| 20 E | 8.13 | 4.13 | 2.30 2.21 | 2.60 2.51 | |

| | | | | | |
|-------------|------|------|--------------|-------------------------|--|
| 21 E | 8.22 | 4.06 | 2.30 | 2.70 2.51 | |
| 22 Q | 8.29 | 4.02 | 2.30 2.23 | 2.51 | Hε 7.08, 6.48 |
| 23 D | 8.38 | 4.51 | 3.20 2.86 | | |
| 24 L | 8.30 | 4.15 | 2.01 1.91 | 1.91 | δCH3 0.97, 0.93 |
| 25 R | 8.10 | 4.09 | 2.01 | 1.91 1.67 | Hδ 3.24 |
| 26 D | 8.28 | 4.56 | 2.97 2.94 | | |
| 27 I | 8.26 | 3.89 | 2.02 | 1.81, 1.22 γCH3 0.99 | Hδ 0.92 |
| 28 G | 8.20 | 3.95 | | | |
| 29 I | 7.96 | 4.11 | 2.01 | 1.64, 1.33 γCH3 0.98 | Hδ 0.91 |
| 30 S | 7.88 | 4.40 | 4.05 3.97 | | |
| 31 D | 7.96 | 4.92 | 3.00 | | |
| 32 P | | 4.31 | 2.44 2.01 | 2.18 2.13 | Hδ 4.03, 3,97 |
| 33 Q | 8.19 | 4.16 | 2.13 | 2.47 2.42 | Hε 7.24, 6.52 |
| 34 H | 8.05 | 4.47 | 3.46 3.42 | | Hδ 7.30 Hε 8.56 |
| 35 R | 8.23 | 4.06 | 1.99 | 1.84 1.70 | Hδ 3.29 Hε 7.20 |
| 36 R | 8.21 | 4.07 | 1.97 1.83 | 1.70 | Hδ 3.22 Hε 7.19 |
| 37 K | 7.84 | 4.16 | 1.99 | 1.64 1.54 | Hδ 1.75 Hε 3.01 |
| 38 L | 7.84 | 4.16 | 1.85 1.73 | 1.84 | δCH3 0.93 |
| 39 L | 8.03 | 4.14 | 1.88 1.63 | 1.82 | δCH3 0.92 |
| 40 Q | 8.01 | 4.08 | 2.24 2.20 | 2.56 2.41 | Hε 7.20, 6.52 |
| 41 A | 8.03 | 4.22 | 1.56 | | |
| 42 A | 8.09 | 4.24 | 1.53 | | |
| 43 R | 7.78 | 4.31 | 2.01 1.90 | 1.82 1.75 | Hδ 3.23 Hε 7.13 CONH ₂ 7.26, 6.77 |

Table 2. Sam3 H_N temperature coefficients ($\Delta\delta/\Delta T$). For Gly7, Leu24, Gly28 and Arg36 residues $\Delta\delta/\Delta T$ values cannot be calculated. The Pro32 is not included in the table.

| Residue | $-\Delta\delta/\Delta T$ (ppb/K) | Residue | $-\Delta\delta/\Delta T$ (ppb/K) |
|----------------|--|----------------|--|
| 1 S | -3.0 ± 0.3 | 22 Q | -5.2 ± 0.3 |
| 2 K | -8.7 ± 0.1 | 23 D | -5.7 ± 0.2 |
| 3 L | -1.2 ± 0.3 | 24 L | --- |
| 4 L | -3.0 ± 0.3 | 25 R | -6.6 ± 0.2 |
| 5 L | -6.6 ± 0.2 | 26 D | -6.5 ± 0.3 |
| 6 N | -5.7 ± 0.2 | 27 I | -8.0 ± 0.3 |
| 7 G | --- | 28 G | --- |
| 8 F | -9.0 ± 0.4 | 29 I | -8.1 ± 0.2 |
| 9 D | -11.9 ± 0.3 | 30 S | -3.4 ± 0.3 |
| 10 D | -9.2 ± 0.2 | 31 D | -3.0 ± 0.2 |
| 11 V | -11.0 ± 0.3 | 33 Q | -1.7 ± 0.2 |
| 12 H | -3.47 ± 0.08 | 34 H | -1.6 ± 0.3 |
| 13 F | -12.2 ± 0.2 | 35 R | -3.2 ± 0.4 |
| 14 L | -13.5 ± 0.4 | 36 R | --- |
| 15 G | -11.9 ± 0.3 | 37 K | -1.6 ± 0.2 |
| 16 S | -5.2 ± 0.3 | 38 L | -4.5 ± 0.3 |
| 17 N | 2.2 ± 0.2 | 39 L | -7.4 ± 0.2 |
| 18 V | -9.0 ± 0.2 | 40 Q | -4.6 ± 0.3 |
| 19 M | -6.1 ± 0.4 | 41 A | -5.4 ± 0.3 |
| 20 E | -5.1 ± 0.2 | 42 A | -7.3 ± 0.3 |
| 21 E | -4.8 ± 0.4 | 43 R | -4.1 ± 0.2 |

ACKNOWLEDGEMENTS

I wish to thank my tutor, Dr. Marilisa Leone, for the great support during this PhD course, and the Institute of Biostructures and Bioimaging (CNR) of Naples. Moreover I thank people who collaborated with us: Dr. Daniela Marasco, Dr. Pasqualina Scognamiglio and Concetta Di Natale, for peptides synthesis, SPR and ITC studies; Dr. Emilia Pedone and Dr. Luciano Pirone for protein expression. Prof. Maurizio Pellecchia from the Sanford|Burnham Medical Research Institute in San Diego is also kindly acknowledged for providing the plasmids used in these studies.

REFERENCES

- Adams R. H., Klein R. Eph receptors and ephrin ligands. essential mediators of vascular development. (2000) *Trends Cardiovasc Med* 10:183-8.
- Altschul S. F., Madden T. L., Schaffer A. A., Zhang J., Zhang Z., Miller W., Lipman D. J. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. (1997) *Nucleic Acids Res* 25:3389-402.
- Aviv T., Lin Z., Lau S., Rendl L. M., Sicheri F., Smibert C. A. The RNA-binding SAM domain of Smaug defines a new family of post-transcriptional regulators. (2003) *Nat Struct Biol* 10:614-21.
- Baron M. K., Boeckers T. M., Vaida B., Faham S., Gingery M., Sawaya M. R., Salyer D., Gundelfinger E. D., Bowie J. U. An architectural framework that may lie at the core of the postsynaptic density. (2006) *Science* 311:531-5.
- Barrera F. N., Poveda J. A., González-Ros J. M., Neira J. L. Binding of the C-terminal sterile alpha motif (SAM) domain of human p73 to lipid membranes. (2003) *J Biol Chem* 278:46878-85.
- Bartels C., Xia T., Billeter M., Güntert P., Wüthrich K. The program XEASY for computer-supported NMR spectral analysis of biological macromolecules. (1995) *J Biomol NMR* 6:1-10.
- Battle E., Bacani J., Begthel H., Jonkheer S., Gregorieff A., van de Born M., Malats N., Sancho E., Boon E., Pawson T., Gallinger S., Pals S., Clevers H. EphB receptor activity suppresses colorectal cancer progression. (2005) *Nature* 435:1126-30.
- Battle E., Henderson J. T., Begthel H., van den Born M. M., Sancho E., Huls G., Meeldijk J., Robertson J., van de Wetering M., Pawson T., Clevers H. Beta-catenin and TCF mediate cell positioning in the intestinal epithelium by controlling the expression of EphB/ephrinB. (2002) *Cell* 111:251-63.
- Baxter N. J., Williamson M. P. Temperature dependence of ¹H chemical shifts in proteins. (1997) *J Biomol NMR* 9:359-69.

Brantley D. M., Cheng N., Thompson E. J., Lin Q., Brekken R. A., Thorpe P. E., Muraoka R. S., Cerretti D. P., Pozzi A., Jackson D., Lin C., Chen J. Soluble Eph A receptors inhibit tumor angiogenesis and progression in vivo. (2002) *Oncogene* 21:7011-26.

Cammers-Goodwin A., Allen T. J., Oslick S. L., McClure K. F., Lee J. H., Kemp D. S. Mechanism of stabilization of helical conformations of polypeptide by water containing trifluoroethanol. (1996) *J Am Chem Soc* 118:3082–90.

Carles-Kinch K., Kilpatrick K. E., Stewart J. C., Kinch M. S. Antibody targeting of the EphA2 tyrosine kinase inhibits malignant cell behavior. (2002) *Cancer Res* 62:2840-7.

Carroll M., Tomasson M. H., Barker G. F., Golub T. R., Gilliland D. G. The TEL/platelet-derived growth factor beta receptor (PDGF beta R) fusion in chronic myelomonocytic leukemia is a transforming protein that self-associates and activates PDGF beta R kinase-dependent signaling pathways. (1996) *Proc Natl Acad Sci U S A* 93:14845-50.

Carter N., Nakamoto T., Hirai H., Hunter T. EphrinA1-induced cytoskeletal re-organization requires FAK and p130. (2002) *Nat Cell Biol* 4:565-73.

Cercone M. A., Schroeder W., Schomberg S., Carpenter T. C. EphA2 receptor mediates increased vascular permeability in lung injury due to viral infection and hypoxia. (2009) *Am J Physiol Lung Cell Mol Physiol* 297: 856–63.

Cheng N., Brantley D. M., Liu H., Lin Q., Enriquez M., Gale N., Yancopoulos G., Cerretti D. P., Daniel T. O., Chen J. Blockade of EphA receptor tyrosine kinase activation inhibits vascular endothelial cell growth factor-induced angiogenesis. (2002) *Mol Cancer Res* 1:2–11.

Cooper M. A., Son A. I., Komlos D., Sun Y., Kleiman N. J., Zhou R. Loss of ephrin-A5 function disrupts lens fiber cell packing and leads to cataract. (2008) *Proc Natl Acad Sci U S A* 105:16620-25.

Coulthard M. G., Morgan M., Woodruff T. M., Arumugam T. V., Taylor S. M., Carpenter T. C., Lackmann M., Boyd A. W. Eph/Ephrin signaling in injury and inflammation. (2012) *Am J Pathol* 181:1493-503

Cowan C. W., Shao Y. R., Sahin M., Shamah S. M., Lin M. Z., Greer P. L., Gao S., Griffith E. C., Brugge J. S., Greenberg M. E. Vav family GEFs link activated Ephs to endocytosis and axon guidance. (2005) *Neuron* 46:205-17.

Craig H. E., Coadwell J., Guillou H., Vermeren S. ARAP3 binding to phosphatidylinositol-(3,4,5)-trisphosphate depends on N-terminal tandem PH domains and adjacent sequences. (2010) *Cell Signal* 22:257-64.

- Dalvit C. Efficient multiple-solvent suppression for the study of the interaction of organic solvents with biomolecules. (1998) *J Biomol NMR* 11: 437-44.
- De Rycker M., Venkatesan R. N., Wei C., Price C. M. Vertebrate tankyrase domain structure and sterile alpha motif (SAM)-mediated multimerization. (2003) *Biochem J* 372:87-96.
- de Vries S. J., van Dijk M., Bonvin A. M. The HADDOCK web server for data-driven biomolecular docking. (2010) *Nat Protoc* 5:883-97.
- Dodelet V. C., Pasquale E. B. Eph receptors and ephrin ligands: embryogenesis to tumorigenesis. (2000) *Oncogene* 19:5614-9.
- Egea, J., Klein, R. Bidirectional Eph-ephrin signaling during axon guidance. (2007). *Trends Cell Biol* 17:230–38.
- Fang W. B., Brantley-Sieders D. M., Parker M. A., Reith A. D., Chen J. A kinase-dependent role for EphA2 receptor in promoting tumor growth and metastasis. (2005) *Oncogene* 24:7859-68.
- Farmer B. T. 2nd, Constantine K. L., Goldfarb V., Friedrichs M. S., Wittekind M., Yanchunas J. Jr., Robertson J. G., Mueller L. Localizing the NADP+ binding site on the MurB enzyme by NMR (1996) *Nat Struct Biol* 3:995-7.
- Farrow N. A., Muhandiram R., Singer A. U., Pascal S. M., Kay C. M., Gish G., Shoelson S. E., Pawson T., Forman-Kay J. D., Kay L. E. Backbone dynamics of a free and phosphopeptide-complexed Src homology 2 domain studied by 15N NMR relaxation. (1994) *Biochemistry* 33:5984-6003.
- Fletcher S., Hamilton A. D. Protein-protein interaction inhibitors: small molecules from screening techniques. (2007) *Curr Top Med Chem* 7:922-7.
- Gale N. W., Holland S. J., Valenzuela D. M., Flenniken A., Pan L., Ryan T. E., Henkemeyer M., Strebhardt K., Hirai H., Wilkinson D. G., Pawson T., Davis S., Yancopoulos G. D. Eph receptors and ligands comprise two major specificity subclasses and are reciprocally compartmentalized during embryogenesis. (1996) *Neuron* 17:9-19.
- Gambardella L., Hemberger M., Hughes B., Zudaire E., Andrews S., Vermeren S. PI3K signaling through the dual GTPase-activating protein ARAP3 is essential for developmental angiogenesis. (2010) *Sci Signal* 3:ra76.
- Gherzi E., Vito P., Lopez P., Abdallah M., D' Adamio L. The intracellular localization of amyloid beta protein precursor (A β PP) intracellular domain associated protein-1 (AIDA-1) is regulated by A β PP and alternative splicing. (2004) *J Alzheimers Dis* 6:67-78.

- Golub T. R., Goga A., Barker G. F., Afar D. E., McLaughlin J., Bohlander S. K., Rowley J. D., Witte O. N., Gilliland D. G. Oligomerization of the ABL tyrosine kinase by the Ets protein TEL in human leukemia. (1996) *Mol Cell Biol* 16:4107-16.
- Green J. B., Gardner C. D., Wharton R. P., Aggarwal A. K. RNA recognition via the SAM domain of Smaug. (2003) *Mol Cell* 11:1537-48.
- Griesinger C., Otting G., Wüthrich K., Ernst R. R. Clean TOCSY for proton spin system identification in macromolecules. (1988) *J Am Chem Soc* 110: 7870-2.
- Groeger G., Nobes C. D. Co-operative Cdc42 and Rho signalling mediates ephrinB-triggered endothelial cell retraction. (2007) *Biochem J* 404:23-9.
- Grzesiek S., Bax A. Amino acid type determination in the sequential assignment procedure of uniformly ¹³C/¹⁵N-enriched proteins. (1993) *J Biomol NMR* 3:185-204.
- Hafner C., Becker B., Landthaler M., Vogt T. Expression profile of Eph receptors and ephrin ligands in human skin and downregulation of EphA1 in nonmelanoma skin cancer. (2006) *Mod Pathol* 19:1369-77.
- Harada B. T., Knight M. J., Imai S., Qiao F., Ramachander R., Sawaya M. R., Gingery M., Sakane F., Bowie J. U. Regulation of enzyme localization by polymerization: polymer formation by the SAM domain of diacylglycerol kinase delta1. (2008) *Structure* 16:380-7.
- Herrmann T., Guntert P., Wüthrich K. Protein NMR structure determination with automated NOE assignment using the new software CANDID and the torsion angle dynamics algorithm DYANA. (2002) *J Mol Biol* 319:209-27.
- Himanen J. P., Chumley M. J., Lackmann M., Li C., Barton W. A., Jeffrey P. D., Vearing C., Geleick D., Feldheim D. A., Boyd A. W., Henkemeyer M., Nikolov D. B. Repelling class discrimination: ephrin-A5 binds to and activates EphB2 receptor signaling. (2004) *Nat Neurosci* 7:501-9.
- Himanen J. P., Saha N., Nikolov D. B. Cell-cell signaling via Eph receptors and ephrins. (2007). *Curr Opin Cell Biol* 19:534–42.
- Huusko P., Ponciano-Jackson D., Wolf M., Kiefer J. A., Azorsa D. O., Tuzmen S., Weaver D., Robbins C., Moses T., Allinen M., Hautaniemi S., Chen Y., Elkahloun A., Basik M., Bova G. S., Bubendorf L., Lugli A., Sauter G., Schleutker J., Ozcelik H., Elowe S., Pawson T., Trent J. M., Carpten J. D., Kallioniemi O. P., Mousset S. Nonsense-mediated decay microarray analysis identifies mutations of EPHB2 in human prostate cancer. (2004) *Nat Genet* 36:979-83.
- I S. T., Nie Z., Stewart A., Najdovska M., Hall N. E., He H., Randazzo P. A., Lock P. ARAP3 is transiently tyrosine phosphorylated in cells attaching to fibronectin and

inhibits cell spreading in a RhoGAP-dependent manner. (2004) *J Cell Sci* 117:6071-84.

Inoue H., Baba T., Sato S., Ohtsuki R., Takemori A., Watanabe T., Tagaya M., Tani K. Roles of SAM and DDHD domains in mammalian intracellular phospholipase A1 KIAA0725p. (2012) *Biochim Biophys Acta* 1823:930-9.

Ireton R. C., Chen J. EphA2 receptor tyrosine kinase as a promising target for cancer therapeutics. (2005) *Curr Cancer Drug Targets* 5:149-57.

Ivanov A. I., Romanovsky A. A. Putative dual role of ephrin-Eph receptor interactions in inflammation. (2006) *IUBMB Life* 58:389-94.

Johnsson B., Lofas S., Lindquist G. Immobilization of proteins to a carboxymethyl-dextran-modified gold surface for biospecific interaction analysis in surface plasmon resonance sensors. (1991) *Anal Biochem* 198:268-77.

Jun G., Guo H., Klein B. E., Klein R., Wang J. J., Mitchell P., Miao H., Lee K. E., Joshi T., Buck M., Chugha P., Bardenstein D., Klein A. P., Bailey-Wilson J. E., Gong X., Spector T. D., Andrew T., Hammond C. J., Elston R. C., Iyengar S. K., Wang B. EPHA2 is associated with age-related cortical cataract in mice and humans. (2009) *PLoS Genet* 5:e1000584.

Kanehisa M., Goto S., Sato Y., Furumichi M., Tanabe M. KEGG for integration and interpretation of large-scale molecular datasets. (2012) *Nucleic Acids Res* 40:109-14.

Kang J. U., Koo S. H., Kwon K. C., Park J. W., Kim J. M. Identification of novel candidate target genes, including EPHB3, MASP1 and SST at 3q26.2-q29 in squamous cell carcinoma of the lung. (2009) *BMC Cancer* 9:237.

Kay L. E., Torchia D. A., Bax A. Backbone dynamics of proteins as studied by 15N inverse detected heteronuclear NMR spectroscopy: application to staphylococcal nuclease. (1989) *Biochemistry* 28:8972-79.

Kim C. A., Bowie J. U. SAM domains: uniform structure, diversity of function. (2003) *Trends Biochem Sci* 28:625-28.

Kim C. A., Gingery M., Pilpa R. M., Bowie J. U. The SAM domain of polyhomeotic forms a helical polymer (2002) *Nat Struct Biol* 9:453-7.

Kim C. A., Phillips M. L., Kim W., Gingery M., Tran H. H., Robinson M. A., Faham S., Bowie J. U. Polymerization of the SAM domain of TEL in leukemogenesis and transcriptional repression. (2001) *EMBO J* 20:4173-82.

Kim C. A., Sawaya M. R., Cascio D., Kim W., Bowie J. U. Structural organization of a Sex-comb-on-midleg/polyhomeotic copolymer. (2005) *J Biol Chem* 280:27769-75.

Kim J., Lee H., Kim Y., Yoo S., Park E., Park S. The SAM domains of Anks family proteins are critically involved in modulating the degradation of EphA receptors. (2010) *Mol Cell Biol* 30:1582-92.

Kitamura T., Kabuyama Y., Kamataki A., Homma M. K., Kobayashi H., Aota S., Kikuchi S., Homma Y. Enhancement of lymphocyte migration and cytokine production by ephrinB1 system in rheumatoid arthritis. (2008) *Am J Physiol Cell Physiol* 294:189–96.

Koolpe M., Dail M., Pasquale E. B. An ephrin mimetic peptide that selectively targets the EphA2 receptor. (2002) *J Biol Chem* 277:46974-79.

Koradi R., Billeter M., Wüthrich K. MOLMOL: a program for display and analysis of macromolecular structures. (1996) *J Mol Graph* 14:51-55.

Krugmann S., Anderson K. E., Ridley S. H., Risso N., McGregor A., Coadwell J., Davidson K., Eguinoa A., Ellson C. D., Lipp P., Manifava M., Ktistakis N., Painter G., Thuring J. W., Cooper M. A., Lim Z. Y., Holmes A. B., Dove S. K., Michell R. H., Grewal A., Nazarian A., Erdjument-Bromage H., Tempst P., Stephens L. R., Hawkins P. T. Identification of ARAP3, a novel PI3K effector regulating both Arf and Rho GTPases, by selective capture on phosphoinositide affinity matrices. (2002) *Mol Cell* 9:95.

Krugmann S., Andrews S., Stephens L., Hawkins P. T. ARAP3 is essential for formation of lamellipodia after growth factor stimulation. (2006) *J Cell Sci* 119:425-32.

Kumar A., Ernst R. R., Wüthrich K. A. Two-dimensional nuclear Overhauser enhancement (2D NOE) experiment for the elucidation of complete proton-proton cross-relaxation networks in biological macromolecules. (1980) *Biochem Biophys Res Commun* 95:1-6.

Kurabi A., Brener S., Mobli M., Kwan J. J., Donaldson L. W. A nuclear localization signal at the SAM-SAM domain interface of AIDA-1 suggests a requirement for domain uncoupling prior to nuclear import. (2009) *J Mol Biol* 392:1168-77.

Labrador J. P., Brambilla R., Klein R. The N-terminal globular domain of Eph receptors is sufficient for ligand binding and receptor signaling. (1997) *EMBO J* 16:3889-97.

Laskowski R. A., Rullmann J. A., MacArthur M. W., Kaptein R., Thornton J. M. AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. (1996) *J Biomol NMR* 8:477-86.

Lee H. J., Hota P. K., Chugha P., Guo H., Miao H., Zhang L., Kim S. J., Stetzk L., Wang B. C., Buck M. NMR structure of a heterodimeric SAM:SAM complex: characterization and manipulation of EphA2 binding reveal new cellular functions of SHIP2. (2011) *Structure* 20:41-55.

- Leone M., Cellitti J., Pellecchia M. NMR studies of a heterotypic Sam-Sam domain association: the interaction between the lipid phosphatase Ship2 and the EphA2 receptor. (2008) *Biochemistry* 47:12721-28.
- Leone M., Cellitti J., Pellecchia M. The Sam domain of the lipid phosphatase Ship2 adopts a common model to interact with Arap3-Sam and EphA2-Sam. (2009) *BMC Struct Biol* 9:59.
- Letunic I., Doerks T., Bork P. SMART 7: recent updates to the protein domain annotation resource. (2012) *Nucleic Acids Res* 40:302-5.
- Li H., Fung K. L., Jin D. Y., Chung S. S., Ching Y. P., Ng I. O., Sze K. H., Ko B. C., Sun H. Solution structures, dynamics, and lipid-binding of the sterile alpha-motif domain of the deleted in liver cancer 2. (2007) *Proteins* 67:1154-66.
- Li J. J., Liu D. P., Liu G. T., Xie D. EphrinA5 acts as a tumor suppressor in glioma by negative regulation of epidermal growth factor receptor. (2009) *Oncogene*. 28:1759-68.
- Lo Conte L., Chothia C., Janin J. The atomic structure of protein-protein recognition sites. (1999) *J Mol Biol* 285:2177-98.
- MacArthur M. W., Thornton J. M. Influence of proline residues on protein conformation. (1991) *J Mol Biol* 218:397-412.
- Marston D. J., Dickinson S., Nobes C. D. Rac-dependent trans-endocytosis of ephrinBs regulates Eph-ephrin contact repulsion. (2003) *Nat Cell Biol* 5:879-88.
- Matsuda S., Harries J. C., Viskaduraki M., Troke P. J., Kindle K. B., Ryan C., Heery D. M. A Conserved alpha-helical motif mediates the binding of diverse nuclear proteins to the SRC1 interaction domain of CBP. (2004) *J Biol Chem* 279:14055-64.
- Mercurio F. A., Marasco D., Pirone L., Pedone E. M., Pellecchia M., Leone M. Solution structure of the first Sam domain of Odin and binding studies with the EphA2 receptor. (2012) *Biochemistry* 51:2136-45.
- Mercurio F. A., Marasco D., Pirone L., Scognamiglio P. L., Pedone E. M., Pellecchia M., Leone M. Heterotypic Sam-Sam association between Odin-Sam1 and Arap3-Sam: binding affinity and structural insights. (2013) *Chembiochem* 14:100-6.
- Meruelo A. D., Bowie J. U. Identifying polymer-forming SAM domains. (2009) *Proteins* 74:1-5.
- Miao H., Burnett E., Kinch M., Simon E., Wang B. Activation of EphA2 kinase suppresses integrin function and causes focal-adhesion-kinase dephosphorylation. (2000) *Nat Cell Biol* 2:62-9.

Morris K., Johnson C. S. Diffusion-ordered two-dimensional nuclear magnetic resonance spectroscopy. (1992) *J Am Chem Soc* 114:3139-41.

Neri D., Szyperski T., Otting G., Senn H., Wüthrich K. Stereospecific nuclear magnetic resonance assignments of the methyl groups of valine and leucine in the DNA-binding domain of the 434 repressor by biosynthetically directed fractional ¹³C labeling, (1989) *Biochemistry* 28:7510-16.

Nieddu E., Pasa S. Interfering with protein-protein contact: molecular interaction maps and peptide modulators. (2007) *Curr Top Med Chem* 7:21-32.

Noberini R., Koolpe M., Peddibhotla S., Dahl R., Su Y., Cosford N. D., Roth G. P., Pasquale E. B. Small molecules can selectively inhibit ephrin binding to the EphA4 and EphA2 receptors. (2008) *J Biol Chem* 283:29461-72.

Noblitt L. W., Bangari D. S., Shukla S., Knapp D. W., Mohammed S., Kinch M. S., Mittal S. K. Decreased tumorigenic potential of EphA2-overexpressing breast cancer cells following treatment with adenoviral vectors that express EphrinA1. (2004) *Cancer Gene Ther* 11:757-66.

Noren N. K., Pasquale E. B. Eph receptor-ephrin bidirectional signals that target Ras and Rho proteins. (2004) *Cell Signal* 16:655-66.

Noren N. K., Pasquale E. B. Paradoxes of the EphB4 receptor in cancer. (2007) *Cancer Res* 67:3994-7.

Notredame C., Higgins D. G., Heringa J. T-Coffee: A novel method for fast and accurate multiple sequence alignment. (2000) *J Mol Biol* 302:205

Pandey A., Blagoev B., Kratchmarova I., Fernandez M., Nielsen M., Kristiansen T. Z., Ohara O., Podtelejnikov A. V., Roche S., Lodish H. F., Mann M. Cloning of a novel phosphotyrosine binding domain containing molecule, Odin, involved in signaling by receptor tyrosine kinases. (2002) *Oncogene* 21:8029-36.

Park J. E., Son A. I., Hua R., Wang L., Zhang X., Zhou R. Human cataract mutations in EPHA2 SAM domain alter receptor stability and function. (2012) *PLoS One* 7:36564.

Parthasarathi L., Casey F., Stein A., Aloy P., Shields D. C. Approved drug mimics of short peptide ligands from protein interaction motifs. (2008) *J Chem Inf Model* 48:1943-8.

Pasquale E. B. Eph-ephrin bidirectional signaling in physiology and disease. (2008) *Cell* 133:38-52.

Pasquale E. B. The Eph family of receptors. (1997) *Curr Opin Cell Biol* 9:608-15.

- Pasquale E. B. Eph receptor signalling casts a wide net on cell behaviour. (2005) *Nat Rev Mol Cell Biol* 6:462–75.
- Pasquale E. B. Eph receptors and ephrins in cancer: bidirectional signalling and beyond. (2010) *Nat Rev Cancer* 10:165-80.
- Pellecchia M. Solution nuclear magnetic resonance spectroscopy techniques for probing intermolecular interactions. (2005) *Chem Biol* 12:961-71.
- Piantini U., Sorensen O. W., Ernst R. R. Multiple quantum filters for elucidating NMR coupling networks. (1982) *J Am Chem Soc* 104:6800-1
- Pitulescu M. E., Adams R. H. Eph/ephrin molecules- a hub for signaling and endocytosis. (2010) *Genes Dev* 24:2480-92.
- Ponting C. P. SAM: A novel motif in yeast sterile and *Drosophila* polyhomeotic proteins. (1995) *Protein Sci* 4:1928-30.
- Price W. S. Pulsed-field gradient nuclear magnetic resonance as a tool for studying translational diffusion: Part 1. Basic theory. (1997) *Concepts Magn Reson* 9:299-306.
- Qiao F., Bowie J. U., The many faces of SAM (2005) *Sci. STKE* 286:re7.
- Qiao F., Song H., Kim C. A., Sawaya M. R., Hunter J. B., Gingery M., Rebay I., Courey A. J., Bowie J. U. Derepression by depolymerization; structural insights into the regulation of Yan by Mae. (2004) *Cell* 118:163-73.
- Raaijmakers J. H., Deneubourg L., Rehmann H., de Koning J., Zhang Z., Krugmann S., Erneux C., Bos J. L. The PI3K effector Arap3 interacts with the PI(3,4,5)P3 phosphatase SHIP2 in a SAM domain-dependent manner. (2007) *Cell Signal* 19:1249-57.
- Rajakulendran T., Sahmi M., Kurinov I., Tyers M., Therrien M., Sicheri F. CNK and HYP form a discrete dimer by their SAM domains to mediate RAF kinase signaling. (2008) *Proc Natl Acad Sci U S A* 105:2836-41.
- Ramachander R., Bowie J. U. SAM domains can utilize similar surfaces for the formation of polymers and closed oligomers. (2004) *J Mol Biol* 342:1353-8.
- Rich R. L., Myszka D. GBIACORE J: a new platform for routine biomolecular interaction analysis. (2001) *J Mol Recognit* 14:223-8.
- Rinne T., Bolat E., Meijer R., Scheffer H., van Bokhoven H. Spectrum of p63 mutations in a selected patient cohort affected with ankyloblepharon–ectodermal defects–cleft lip/palate syndrome (AEC). (2009) *Am J Med Genet A* 149:1948–51.

- Roccatano D., Colombo G., Fioroni M., Mark A. E. Mechanism by which 2,2,2-trifluoroethanol/water mixtures stabilize secondary-structure formation in peptides: a molecular dynamics study. (2002) *Proc Natl Acad Sci U S A* 99:12179-84.
- Ruoslahti E. Fibronectin and its integrin receptors in cancer. (1999) *Adv Cancer Res* 76:1-20.
- Schultz J., Milpetz F., Bork P., Ponting C. P. SMART, a simple modular architecture research tool: Identification of signaling domains. (1998) *Proc Natl Acad Sci USA* 95:5857-64.
- Schultz J., Ponting C. P., Hofmann K., Bork P. SAM as a protein interaction domain involved in developmental regulation. (1996) *Protein Sci* 6:249-53.
- Shiels A., Bennett T. M., Knopf H. L., Maraini G., Li A., Jiao X., Hejtmancik J. F. The EPHA2 gene is associated with cataracts linked to chromosome 1p. (2008) *Mol Vis* 14:2042-55.
- Smalla M., Schmieder P., Kelly M., Ter Laak A., Krause G., Ball L., Wahl M., Bork P., Oschkinat H. Solution structure of the receptor tyrosine kinase EphB2 SAM domain and identification of two distinct homotypic interaction sites. (1999) *Protein Sci* 8:1954-61.
- Stafford R. L., Hinde E., Knight M. J., Pennella M. A., Ear J., Digman M. A., Gratton E., Bowie J. U. Tandem SAM domain structure of human Caskin1: a presynaptic, self-assembling scaffold for CASK. (2011) *Structure* 19:1826-36.
- Stapleton D., Balan I., Pawson T., Sicheri F. The crystal structure of an Eph receptor SAM domain reveals a mechanism for modular dimerization. (1999) *Nat Struct Biol* 6:44-9.
- Sulman E. P., Tang X. X., Allen C., Biegel J. A., Pleasure D. E., Brodeur G. M., Ikegaki N. ECK, a human EPH-related gene, maps to 1p36.1, a common region of alteration in human cancers. (1997) *Genomics* 40:371-4.
- Surawska H., Ma P. C., Salgia R. The role of ephrins and Eph receptors in cancer. (2004) *Cytokine Growth Factor Rev* 15:419-33.
- Talluri S., Wagner G. An optimized 3D NOESY-HSQC. (1996) *J Magn Reson B* 112:200-5.
- Thanos C. D., Faham S., Goodwill K. E., Cascio D., Phillips M., Bowie J. U. Monomeric structure of the human EphB2 sterile alpha motif domain. (1999) *J Biol Chem* 274:37301-6.
- Thanos C. D., Goodwill K. E., Bowie J. U. Oligomeric structure of the human EphB2 receptor SAM domain. (1999) *Science* 283:833-6.

Tognolini M., Incerti M., Hassan-Mohamed I., Giorgio C., Russo S., Bruni R., Lelli B., Bracci L., Noberini R., Pasquale E. B., Barocelli E., Vicini P., Mor M., Lodola A. Structure-activity relationships and mechanism of action of Eph-ephrin antagonists: interaction of cholanic acid with the EphA2 receptor. (2012) *Chem Med Chem* 7:1071-83.

van Dijk A. D., Bonvin A. M. Solvated docking: introducing water into the modelling of biomolecular complexes. (2006) *Bioinformatics* 22:2340-7.

Vanhaesebroeck B., Leever S. J., Panayotou G., Waterfield M. D. Phosphoinositide 3-kinases: a conserved family of signal transducers. (1997) *Trends Biochem Sci* 22:267-72.

Wei Z., Zheng S., Spangler S. A., Yu C., Hoogenraad C. C., Zhang M. Liprin-mediated large signaling complex organization revealed by the liprin- α /CASK and liprin- α /liprin- β complex structures. (2011) *Mol Cell* 43:586-98.

White T. E., Brandariz-Nuñez A., Carlos Valle-Casuso J., Amie S., Nguyen L., Kim B., Brojatsch J., Diaz-Griffero F. Contribution of SAM and HD domains to retroviral restriction mediated by human SAMHD1. (2012) *Virology* 436:81-90.

Wilkin D. K., Grimshaw S. B., Receveur V., Dobson C. M., Jones J. A., Smith L. J. Hydrodynamic radii of native and denatured proteins measured by pulse field gradient NMR techniques. (1999) *Biochemistry* 38:16424-31.

Wilkinson D. G. Multiple roles of EPH receptors and ephrins in neural development. (2001) *Nat Rev Neurosci* 2:155-64.

Winter J., Roepcke S., Krause S., Müller E. C., Otto A., Vingron M., Schweiger S. Comparative 3'UTR analysis allows identification of regulatory clusters that drive Eph/ephrin expression in cancer cell lines. (2008) *PLoS One* 3:e2780.

Wishart D. S., Sykes B. D., Richards F. M. Relationship between nuclear magnetic resonance chemical shift and protein secondary structure. (1991) *J Mol Biol* 222:311-33.

Wüthrich K. *NMR of proteins and nucleic acids*. (1986) Wiley, New York.

Yagi R., Tanaka M., Sasaki K., Kamata R., Nakanishi Y., Kanai Y., Sakai R. ARAP3 inhibits peritoneal dissemination of scirrhous gastric carcinoma cells by regulating cell adhesion and invasion. (2011) *Oncogene* 30:1413-21.

Zelinski D. P., Zantek N. D., Stewart J. C., Irizarry A. R., Kinch M. S. EphA2 overexpression causes tumorigenesis of mammary epithelial cells. (2001) *Cancer Res* 61:2301-6.

Zhang H., Xu Q., Krajewski S., Krajewska M., Xie Z., Fuess S., Kitada S., Pawlowski K., Godzik A., Reed J. C. BAR: An apoptosis regulator at the intersection of caspases and Bcl-2 family proteins. (2000) *Proc Natl Acad Sci U S A* 97:2597-602.

Zhang T., Hua R., Xiao W., Burdon K. P., Bhattacharya S. S., Craig J. E., Shang D., Zhao X., Mackey D. A., Moore A. T., Luo Y., Zhang J., Zhang X. Mutations of the EPHA2 receptor tyrosine kinase gene cause autosomal dominant congenital cataract. (2009) *Hum Mutat* 30:603-11.

Zhuang G., Brantley-Sieders D. M., Vaught D., Yu J., Xie L., Wells S., Jackson D., Muraoka-Cook R., Arteaga C., Chen J. Elevation of receptor tyrosine kinase EphA2 mediates resistance to trastuzumab therapy. (2010) *Cancer Res* 70:299-308.

Zhuang G., Hunter S., Hwang Y., Chen J. Regulation of EphA2 receptor endocytosis by SHIP2 lipid phosphatase via phosphatidylinositol 3-Kinase-dependent Rac1 activation. (2007) *J Biol Chem* 282:2683-94.

Zimmer M., Palmer A., Köhler J., Klein R. EphB-ephrinB bi-directional endocytosis terminates adhesion allowing contact mediated repulsion. (2003) *Nat Cell Biol* 5:869-78.

LIST OF PUBLICATIONS

Flavia A. Mercurio, Daniela Marasco, Luciano Pirone, Emilia M. Pedone, Maurizio Pellecchia & Marilisa Leone . Solution structure of the first Sam domain of Odin and binding studies with the EphA2 receptor. *Biochemistry*. 2012;51(10):2136-45.- Selected by F1000.

Flavia A. Mercurio, Daniela Marasco, Luciano Pirone, Pasqualina L. Scognamiglio, Emilia M. Pedone, Maurizio Pellecchia, & Marilisa Leone. Heterotypic Sam-Sam association between Odin-Sam1 and Arap3-Sam: binding affinity and structural insights. *Chembiochem*. 2013 (14):100-106.

Flavia A. Mercurio, Susan Costantini, Raffaele Raucci, Ringhieri Paola, Marilisa Leone, Filomena Rossi, Giovanni Colonna, Giuseppe Castello, & Stefania Scala. Short and flexible peptides targeting chemokine receptor CXCR4. *Chemokine* (In preparation).

LIST OF POSTER PRESENTATIONS

September 2012, XLI National Congress on Magnetic Resonance, Pisa.
Title: “ NMR studies of the first Sam domain of Odin and its association with the Sam domain of EphA2”
F. A. Mercurio, D. Marasco, L. Pirone, E. M. Pedone, M. Pellecchia & M. Leone

June 2012, 13th Naples Workshop on Bioactive Peptides, Naples.
Title: “Sam domains heterotypic interactions: functional and structural characterization of minimum interacting protein regions”
C. Di Natale, F. Mercurio, P.L. Scognamiglio, L. Pirone, E. M. Pedone, M. Pellecchia, D. Marasco, & M. Leone.

September 2011, Artzymes, Marseille.
Title: “NMR studies of heterotypic Sam-Sam domain interactions involving EphA2 receptor”
F. A. Mercurio, D. Marasco, L. Pirone, E. M. Pedone, E. Benedetti, M. Pellecchia & M. Leone.

April 2011, 52nd Experimental Nuclear Magnetic Resonance Conference, Asilomar, California.
Title: “NMR studies of the heterotypic Sam-Sam domain interaction between Odin and the EphA2 receptor”
F. A. Mercurio, D. Marasco, L. Pirone, E. M. Pedone, M. Pellecchia & M. Leone.

ORAL COMMUNICATIONS

September 2011, XL National Congress on Magnetic Resonance, Parma.
Title: “NMR studies of the heterotypic Sam-Sam domain associations involving EphA2 receptor”
F. A. Mercurio, D. Marasco, L. Pirone, E. M. Pedone, M. Pellecchia & M. Leone.

ATTENDED COURSES AND SCHOOLS

21/01/2013-25/01/2013

“Advanced methods for the integration of other structural data with NMR data”,
Magnetic Resonance Center (CERM), University of Florence

13/04/2012-30/11/2012

English course for PhD students- University Language Centre of University of Naples
“Federico II”. First Cambridge Certificate (FCE) obtained (Level B2 of the Common
European Framework of Reference).

12/12/2011-14/12/2011

Intensive theoretical and experimental course on protein crystallization using manual
and / or robotic techniques, Institute of Biostructures and Bioimaging of National
Research Council (CNR) of Naples.

28/11/2011-29/11/2011

“Young professionals in Life Sciences: opportunities and challenges in convergence
regions”, Molecular Diagnostics and Pharmaceutical (DFM Scarl), Naples.

30/08/2010-03/09/2010

Nuclear Magnetic Resonance School “Structure, dynamics and interactions of
proteins”- Italian NMR Discussion Group (GIDRM), Turin.

Solution Structure of the First Sam Domain of Odin and Binding Studies with the EphA2 Receptor

Flavia Anna Mercurio,[†] Daniela Marasco,^{§,†,‡} Luciano Pirone,[§] Emilia Maria Pedone,^{§,‡} Maurizio Pellicchia,^{||} and Marilisa Leone^{*,§,‡}

[†]Department of Biological Sciences, University of Naples “Federico II”, Naples, Italy

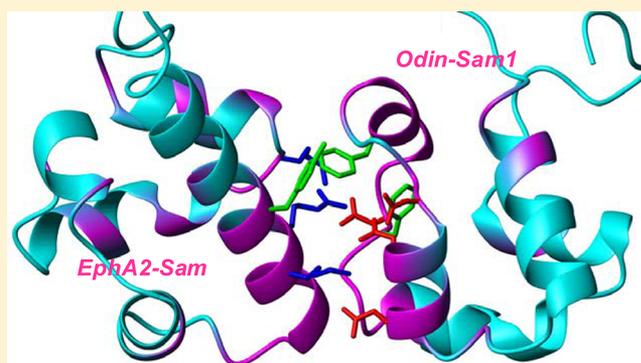
[‡]Centro Interuniversitario di Ricerca sui Peptidi Bioattivi (CIRPEB), Naples, Italy

[§]Institute of Biostructures and Bioimaging, National Research Council, Naples, Italy

^{||}Sanford-Burnham Medical Research Institute, La Jolla, California 92037, United States

S Supporting Information

ABSTRACT: The EphA2 receptor plays key roles in many physiological and pathological events, including cancer. The process of receptor endocytosis and the consequent degradation have attracted attention as possible means of overcoming the negative outcomes of EphA2 in cancer cells and decreasing tumor malignancy. A recent study indicates that Sam (sterile alpha motif) domains of Odin, a member of the ANKS (ankyrin repeat and sterile alpha motif domain-containing) family of proteins, are important for the regulation of EphA2 endocytosis. Odin contains two tandem Sam domains (Odin-Sam1 and -Sam2). Herein, we report on the nuclear magnetic resonance (NMR) solution structure of Odin-Sam1; through a variety of assays (employing NMR, surface plasmon resonance, and isothermal titration calorimetry techniques), we clearly demonstrate that Odin-Sam1 binds to the Sam domain of EphA2 in the low micromolar range. NMR chemical shift perturbation experiments and molecular modeling studies point out that the two Sam domains interact with a head-to-tail topology characteristic of several Sam–Sam complexes. This binding mode is similar to that we have previously proposed for the association between the Sam domains of the lipid phosphatase Ship2 and EphA2. This work further validates structural elements relevant for the heterotypic Sam–Sam interactions of EphA2 and provides novel insights for the design of potential therapeutic compounds that can modulate receptor endocytosis.



Eph receptors represent a large subgroup of the receptor tyrosine kinase family and together with their ephrin ligands play relevant roles in several physiological and pathological processes.^{1,2} Interestingly, these receptors are differentially expressed in unhealthy versus normal tissues and, thus, considered attractive targets in drug discovery.³ Among them, EphA2 has received a great amount of attention.

The most recent work has associated EphA2 with cataracts^{4–6} and entry of the hepatitis C virus into the host cell.⁷ However, EphA2 has long been related to cancer, and even if its role in this disease has been described as both complex and controversial, many of its procancer activities have been well characterized.⁴

The processes of enhanced EphA2 endocytosis and subsequent degradation have been correlated with weakened malignant cell behavior.¹ Recent studies have focused on the regulation mechanisms at the basis of EphA2 receptor endocytosis.^{8,9} First, lipid phosphatase Ship2 (Src homology 2 domain-containing phosphoinositide-5-phosphatase 2) has been identified as a prominent regulator of this process.⁸ In vitro experiments have demonstrated that Ship2 overexpression

in malignant breast cancer cells increases EphA2 stability, while decreased levels of Ship2 facilitate receptor internalization and degradation. To exert its function, Ship2 needs to be engaged at the receptor site by means of a heterotypic interaction between its sterile alpha motif (Sam) domain and the Sam domain of EphA2 (EphA2-Sam).⁸

Sam domains are protein binding modules containing ~70 amino acids that form a five-helix bundle and are involved in many biological processes mainly via homo- and heterodimerization or polymerization processes.^{10,11}

The nuclear magnetic resonance (NMR) solution structure of the Sam domain of Ship2 [Ship2-Sam, Protein Data Bank (PDB) entry 2K4P] and binding studies with the Sam domain from the EphA2 receptor (EphA2-Sam) have previously been reported.¹² ITC (isothermal titration calorimetry) experiments have indicated that the two domains interact with a dissociation constant in the low micromolar range. Furthermore, chemical shift perturbation experiments have revealed

Received: January 31, 2012

Published: February 14, 2012

the reciprocal binding interfaces of Ship2-Sam and EphA2-Sam and, together with modeling studies, have shown that this interaction may adopt a mid loop (ML)/end helix (EH) model, characteristic of other Sam-Sam complexes.^{12–15}

A recent work indicates that EphA2 endocytosis is also regulated by ANKS (ankyrin repeat and Sam domain-containing) family proteins.⁹ This class of proteins includes Odin¹⁶ and AIDA1b (A β PP intracellular domain-associated protein 1B),¹⁷ which possess in addition to six ankyrin repeats and a PTB (phosphotyrosine binding) domain, two Sam domains in tandem (Sam1 and Sam2).

Overexpression of Odin in MDA-MB-231 human breast carcinoma cells and MEFs (mouse embryonic fibroblasts) protects EphA2 from undertaking internalization and degradation after ligand stimulation, while a Sam domain deletion mutant of Odin lacks this function.⁹

Herein, we describe solution structure studies of the first Sam domain of Odin (Odin-Sam1) and binding studies with EphA2-Sam. Through a variety of assays relying on NMR, surface plasmon resonance (SPR), and ITC techniques, we clearly demonstrate that Odin-Sam1 and EphA2-Sam interact with low micromolar affinity and a 1:1 stoichiometry. NMR chemical shift perturbation experiments allow identification of the reciprocal binding interfaces of the two proteins; moreover, NMR-based displacement experiments and molecular docking studies show that Ship2-Sam and Odin-Sam1 share a common binding site on the surface of EphA2-Sam and adopt similar binding modes for these heterotypic Sam-Sam interactions.

Our work sheds additional light on the structural features that are relevant for heterotypic Sam-Sam complexes involving EphA2 and provides novel information for the design of therapeutic compounds that can modulate receptor endocytosis.

MATERIALS AND METHODS

Protein Expression. Recombinant proteins were expressed in *Escherichia coli*. The following PET15B constructs were used for this study: Ship2-Sam (residues 1199–1258 of human Ship2, UniprotKB/TrEMBL entry O15357), Odin-Sam1 (residues 691–770 of human Odin, UniprotKB/TrEMBL entry Q92625), and Odin-Sam2 (residues 761–840 of human Odin, UniprotKB/TrEMBL entry Q92625). Constructs designed for wild-type human EphA2-Sam (Swiss-Prot/TrEMBL entry P29317) and its double (H45N/R71A) and triple (K38A/R78A/Y81S) mutants have been previously described.^{12,18}

Synthetic genes encoding these proteins were all purchased from Celtek Bioscience (Nashville, TN) and contain an N-terminal His tag and a thrombin cleavage site.

Genes were transformed using BL21-Gold (DE3) competent cells (Stratagene). Protein expression and purification protocols implemented for labeled and unlabeled Sam domain production have previously been reported.¹⁸

To express unlabeled proteins, bacteria were grown in LB medium. M9 minimal medium supplemented with 2 g/L [¹³C]glucose and/or 0.5 g/L ¹⁵NH₄Cl was prepared for expression of ¹⁵N- and ¹³C-labeled or ¹⁵N-labeled proteins, respectively. M9 medium containing 3.6 g/L [¹²C]glucose (natural abundance) and 0.4 g/L [¹³C]glucose was used to achieve 10% fractional ¹³C labeling for stereospecific assignments of Leu CH₃ ^{δ 1,2} and Val CH₃ ^{γ 1,2} groups.¹⁹ Expression protocols for uniformly or selectively labeled proteins were identical to those followed for unlabeled protein production. Briefly, bacteria were grown at 37 °C; protein overexpression was

induced at a cell optical density OD₆₀₀ of 0.6 nm with isopropyl β -D-thiogalactopyranoside (IPTG, 1 mM) at 25 °C overnight.

Proteins were purified with an AKTA Purifier FPLC system by affinity chromatography on a nickel column (GE Healthcare, Milan, Italy).

Resonance Assignments. Resonance assignment experiments were conducted at 25 °C on a Varian Unity Inova 600 MHz spectrometer equipped with a cold probe. NMR samples consisted of ¹⁵N-labeled or ¹⁵N- and ¹³C-labeled Odin-Sam1 (~900 μ M) in phosphate-buffered saline (PBS) (10 mM phosphates, 140 mM NaCl, and 2.7 mM KCl) (Fisher) at pH 7.7 and 0.2% NaN₃ with volumes of 600 μ L (95% H₂O/5% D₂O).

Backbone assignments were made via analysis of triple-resonance experiments [HNCA, HN(CO)CACB, and HNCACB].²⁰ Carbon side chains were identified in (H)CC-(CO)NH and HCCH-TOCSY spectra. Proton side chains were assigned in the HCCH-TOCSY spectrum or by comparing three-dimensional (3D) ¹⁵N-resolved ¹H–¹H NOESY (100 ms mixing time) and 3D ¹⁵N-resolved ¹H–¹H TOCSY (70 ms mixing time) spectra. Aromatic side chains were identified by combined analysis of two-dimensional (2D) ¹H–¹H NOESY (100 ms mixing time) and 2D TOCSY (70 ms mixing time) spectra recorded for samples of Odin-Sam1 dissolved in 99% D₂O. Stereospecific assignments for Leu CH₃ ^{δ 1,2} and Val CH₃ ^{γ 1,2} groups of Odin-Sam1 were obtained from a ¹H–¹³C HSQC experiment of a fractionally ¹³C-labeled Odin-Sam1 sample (500 μ M).¹⁹

To obtain the nearly complete backbone resonance assignments of the Odin-Sam1–EphA2-Sam complex, HNCA and 3D ¹⁵N-resolved ¹H–¹H NOESY (100 ms mixing time) spectra were acquired with samples containing either ¹⁵N- and ¹³C-labeled Odin-Sam1 (450 μ M) and unlabeled EphA2-Sam (~1 mM) or doubly labeled EphA2-Sam (350 μ M) and unlabeled Odin-Sam1 (900 μ M).

NMR spectra were processed with Varian software (Vnmrj version 1.1D) and analyzed with NEASY²¹ (<http://www.nmr.ch/>).

Relaxation Measurements. Backbone ¹⁵N longitudinal (R_1) and transverse (R_2) relaxation rates were evaluated at 25 °C and 600 MHz. Measurements were taken with two ¹⁵N-labeled Odin-Sam1 samples at concentrations of 100 and 900 μ M and a sample of the Sam-Sam complex consisting of labeled Odin-Sam1 (450 μ M) and unlabeled EphA2-Sam (~1 mM).

R_1 and R_2 relaxation data were collected as one-dimensional spectra (4K data points and 2K or 4K transients). R_1 data sets were recorded with relaxation delays of 0.01, 0.1, 0.3, 0.6, and 1.0 s; R_2 data sets were acquired with relaxation delays of 0.01, 0.03, 0.05, 0.07, 0.09, 0.11, 0.15, and 0.19 s. Average R_1 and R_2 values were estimated by the decrease in signal intensity as function of relaxation delay. To calculate the rotational correlation time, we used average R_2/R_1 ratios as input for the software tmest (A. G. Palmer, III, Columbia University, New York, NY).²²

DOSY (Diffusion-Ordered Spectroscopy). Diffusion-ordered NMR spectroscopy was conducted with the pulsed gradient spin-echo (PGSE) NMR technique.²³ The translational self-diffusion coefficient D can be calculated with the equation $I = I_0 \exp[-D\gamma^2\delta^2G^2(\Delta - \delta/3)]$, where I_0 is the measured signal intensity of a set of resonances at the smaller gradient strength, I is the corresponding observed peak intensity, D is the diffusion constant, γ is the proton gyromagnetic ratio, δ is

the diffusion gradient length, G is the gradient strength, and Δ is the diffusion delay.²⁴

Series of spectra were acquired with 512 scans and 16K data points. We conducted DOSY experiments with Odin-Sam1 samples at concentrations of 100 and 900 μM . The hydrodynamic radius (r_{H}) of the protein was evaluated with the Stokes–Einstein equation: $D_{\text{t}} = k_{\text{B}}T/f$, where $f (=6\pi\eta r_{\text{H}})$ is the translational friction coefficient, η is the viscosity of the solution, k_{B} is Boltzmann's constant, and T is the temperature in kelvin.

Structure Calculations and Analysis for Odin-Sam1.

Analysis of a 3D ^{15}N -resolved ^1H – ^1H NOESY-HSQC spectrum²⁵ (100 ms mixing time), a 3D ^{13}C -resolved ^1H – ^1H NOESY-HSQC spectrum (100 ms mixing time), and a 2D ^1H – ^1H NOESY spectrum²⁶ (100 ms mixing time), for the aliphatic to aromatic region, that was acquired after dissolving the lyophilized protein sample in 99% D_2O , was used to collect distance constraints for structure calculations. CYANA version 2.1²⁷ was employed to calculate the solution structure of Odin-Sam1. Angular constraints were generated with the GRID-SEARCH module of CYANA. The final structure calculation includes 1206 upper distance constraints (393 intraresidue, 239 short-range, 275 medium-range, and 299 long-range), 372 angle constraints, and information about stereospecific assignments for methyl groups of Val32, Val56, Val63, Leu36, Leu48, and Leu84. Structure calculations were initiated from 100 random conformers; the 20 structures that better satisfy experimental constraints (i.e., lowest CYANA target functions) were further inspected with MOLMOL²⁸ and iCING (<http://proteins.dyn dns.org/cing/iCing.html>). Surface representations were generated with PVM 1.5.6rc1.²⁹

NMR Binding Studies. Protein–protein interaction studies were conducted by means of NMR titration experiments and analysis of 2D ^1H – ^{15}N HSQC spectra. To identify the Odin-Sam1 binding interface for EphA2-Sam, 2D ^1H – ^{15}N HSQC spectra of ^{15}N -labeled Odin-Sam1 (160 μM) were recorded for the protein in the unbound form and after addition of unlabeled EphA2-Sam (80, 160, 240, and 400 μM). To recognize the binding site of EphA2-Sam for Odin-Sam1, 2D ^1H – ^{15}N HSQC spectra of a ^{15}N -labeled EphA2-Sam sample (63 μM) were recorded in the absence and presence of unlabeled Odin-Sam1 (360 and 600 μM).

To verify the binding of EphA2-Sam mutants to Odin-Sam1, NMR chemical shift perturbation experiments were performed with ^{15}N -labeled Odin-Sam1 samples (50–100 μM) and unlabeled double (H45N/R71A) and triple (K38A/R78A/Y81S) EphA2-Sam mutants (concentrations ranging from 200 μM to 2 mM) (see the Supporting Information).

Analysis of titration experiments and overlays of 2D spectra were performed with SPARKY 3 (T. D. Goddard and D. G. Kneller, University of California, San Francisco).

Surface Plasmon Resonance. EphA2-Sam proteins were immobilized in 10 mM acetate buffer (pH 5.0, flow rate of 5 $\mu\text{L}/\text{min}$, injection time of 7 min) on a CMS Biacore sensor chip, using EDC/NHS chemistry, following the manufacturer's instructions.³⁰ Residual reactive groups were deactivated with 1 M ethanolamine hydrochloride (pH 8.5); the reference channel was prepared by activation with EDC/NHS and deactivation with ethanolamine. Immobilization levels were 940, 1313, and 1480 resonance units (RU) for wild-type EphA2-Sam and double (H45N/R71A) and triple (K38A/R78A/Y81S) mutants, respectively. Experiments were conducted at 25 $^{\circ}\text{C}$ and a constant flow rate of 20 $\mu\text{L}/\text{min}$ using as running

buffer a solution of 10 mM Hepes (pH 7.4), 150 mM NaCl, and surfactant P20 (0.05%, v/v, 90 μL injected for each experiment). Binding experiments were conducted with Odin-Sam1 at various concentrations in the range of 0.1–400 μM . The BIA evaluation analysis package (version 4.1, GE Healthcare) was used to subtract the signal of the reference channel and to estimate K_{D} values. RU_{max} values as a function of protein concentration were fit by nonlinear regression analysis with GraphPad Prism, version 4.00 (GraphPad Software, San Diego, CA).³¹

Isothermal Titration Calorimetry (ITC). ITC studies were performed at 25 $^{\circ}\text{C}$ with an iTC200 calorimeter (MicroCal/GE Healthcare, Milan, Italy). A solution of EphA2-Sam at a concentration of 257 μM was titrated into a solution of Odin-Sam1 (10 μM). Both proteins were extensively dialyzed in the same buffer (PBS, pH 7.7) prior to ITC measurements. Fitting of data to a single-binding site model was conducted with the Origin software as supplied by GE Healthcare. ITC runs were repeated twice to evaluate the reproducibility of the results.

Docking Studies. Models of the Odin-Sam1–EphA2-Sam complex were generated with the Haddock web server.³² NMR structures (first conformers) of both Odin-Sam1 [Protein Data Bank (PDB) entry 2LMR] and EphA2-Sam (PDB entry 2E8N, RIKEN Structural Genomics Initiative) were used in these studies. Ambiguous interaction restraints were generated from chemical shift perturbation data. For Odin-Sam1, active residues (i.e., L49, L50, N51, F53, D54, D55, V56, H57, F58, Q67, D68, R70, and D71) were chosen among those with the highest chemical shift variations because they either possess high solvent exposure or could potentially supply important intermolecular contacts as revealed by structural homology with other Sam–Sam complexes. For EphA2-Sam, we adopted similar selection criteria for active residues (i.e., K38, R71, G74, H75, K77, R78, and Y81). The C-terminal and N-terminal tails of Odin-Sam1 (residues 21–24 and 95–101, respectively) were considered fully flexible during all the docking stages, whereas the region encompassing residues 51–65 was set as semiflexible. For EphA2-Sam, fully flexible segments include the N-terminal and C-terminal tails; the portion of C-terminal $\alpha 5$ helix covering residues 74–84 of EphA2-Sam (corresponding to amino acids 59–69 according to the sequence numbers of PDB entry 2E8N) was considered semiflexible.

In all of the docking runs, passive residues were set automatically by the Haddock web server and the solvated docking mode was turned on.³³

In the first iteration of the docking protocol (i.e., the rigid body energy minimization), 1000 structures were calculated; in the second iteration, the 200 best solutions were subjected to semiflexible simulated annealing, and a final refinement in water was also performed.

The final 200 Haddock models were all visually inspected, and solutions not compatible with our experimental data and/or containing highly unusual protein orientations were soon removed. This first selection screening reduced the number of structures to 78. The resultant solutions were further analyzed with MOLMOL²⁸ and compared with experimental structures of other heterotypic Sam–Sam complexes (for example, the AIDA1 Sam tandem, PDB entry 2KIV³⁴), to recognize characteristic features. At the end of this analysis, 22 models were chosen as being representative of the possible EphA2-Sam–Odin-Sam1 conformations. Solutions were clustered using a pairwise root-mean-square deviation (rmsd) cutoff of 2 \AA and at least one structure per cluster. The rmsd values were

calculated with MOLMOL²⁸ by superimposing the structures on the backbone atoms of residues 36–41, 62–65, and 70–81 of EphA2-Sam and residues 48–73 of Odin-Sam1. This clustering procedure indicated the presence of five families of structures in the selected ensemble (see Figure S3 of the Supporting Information).

RESULTS

NMR Solution Structure of Odin-Sam1. To assess the quaternary structure of Odin-Sam1 in solution, we have conducted ¹⁵N *R*₁ and *R*₂ nuclear spin relaxation rate measurements along with DOSY (diffusion-ordered spectroscopy) experiments.

The rotational correlation time, τ_c , of the protein at 100 μ M, estimated by relaxation data (*R*₂/*R*₁ average value), is 7.4 ± 0.7 ns and does not change when the protein concentration is increased to 900 μ M (i.e., 7.3 ± 0.6 ns). The τ_c of Odin-Sam1 bound to EphA2-Sam increases instead to 11 ± 1 ns.

DOSY measurements²³ indicate for a more diluted Odin-Sam1 sample (100 μ M) a diffusion coefficient (*D*_t) equal to $(1.5 \pm 0.1) \times 10^{-10}$ m² s⁻¹ and a corresponding hydrodynamic radius (*r*_H) of 16.7 Å. For an Odin-Sam1 sample at a higher concentration (900 μ M), we obtained similar diffusion parameters [*D*_t = $(1.40 \pm 0.07) \times 10^{-10}$ m² s⁻¹, and *r*_H = 17.5 Å].

These studies clearly demonstrate that under the experimental conditions used to calculate Odin-Sam1 NMR structure (~900 μ M) aggregation processes can be ignored.

The Odin-Sam1 solution structure is a canonical Sam domain helix bundle (Figure 1), and the region encompassing the

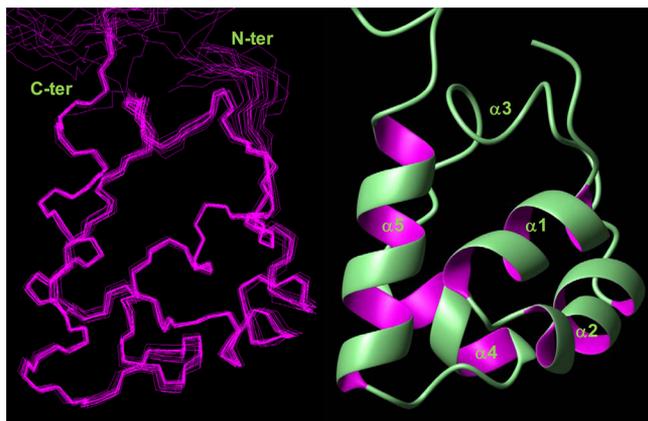


Figure 1. Superposition of the backbone atoms (residues 30–90) of Odin-Sam1 NMR structures (left). Odin-Sam1, first conformer, is shown in ribbon representation (right), including the following α -helical segments: $\alpha 1$ (residues 32–39), $\alpha 2$ (residues 42–50), $\alpha 4$ (residues 66–72), and $\alpha 5$ (residues 77–88). The disordered N-terminal tail encompassing residues 21–27 has been omitted for the sake of clarity.

$\alpha 3$ helix lacks ordered secondary structure elements; however, a reduced number of constraints could be collected for residues in this portion of the protein.

Relevant structural parameters for Odin-Sam1 conformers are listed in Table 1.

Odin-Sam1–EphA2-Sam Interaction. Binding of Odin-Sam1 to EphA2-Sam was first monitored by means of chemical shift perturbation studies.³⁵ 2D ¹H–¹⁵N HSQC experiments were recorded for a uniformly ¹⁵N-labeled Odin-Sam1 sample

Table 1. Statistics for the Odin-Sam1 Solution Structure

| | |
|--|------------------|
| no. of NOE upper distance limits | 1206 |
| no. of angle constraints | 372 |
| residual target function (Å ²) | 0.79 ± 0.08 |
| no. of residual NOE violations | |
| no. >0.1 Å ^a | 2 |
| maximum (Å) | 0.205 ± 0.07 |
| no. of residual angle violations | 0 |
| atomic pairwise rmsd (Å) | |
| backbone atoms (amino acids 30–90) | 0.26 ± 0.07 |
| heavy atoms (amino acids 30–90) | 0.74 ± 0.09 |
| Procheck analysis ^b (%) | |
| residues in core regions | 84.6 |
| residues in allowed regions | 15.1 |
| residues in generous regions | 0.3 |
| residues in disallowed regions | 0 |

^aMaximal number of CYANA²⁷ violations. ^bPROCHECK_NMR⁴⁶ statistics for residues 30–90.

in the presence and absence of unlabeled EphA2-Sam (Figure 2A). Proton- and nitrogen-normalized chemical shift variations were evaluated with the equation $\Delta\delta = [(\Delta H_N)^2 + (0.17 \times \Delta^{15}N)^2]^{1/2}$ (Figure 2B).³⁶ Upon heterotypic association, many changes affect the spectrum; however, the greatest $\Delta\delta$ values (>0.2 ppm) occur in the middle region of the protein, including helices $\alpha 3$, $\alpha 4$, and to a lesser extent $\alpha 2$ (Figure 2C,D).

Similar NMR experiments with ¹⁵N-labeled EphA2-Sam and unlabeled Odin-Sam1 were conducted to map the binding surface of EphA2-Sam for Odin-Sam1 (Figure 3A). We estimated normalized chemical shift deviations to identify the residues of the receptor participating in the interaction with Odin-Sam1 (Figure 3B). The largest deviations are localized at the interface between helix $\alpha 5$ and the adjacent $\alpha 1$ – $\alpha 2$ and $\alpha 4$ – $\alpha 5$ loop regions (Figure 3C).

Chemical shift mapping data suggest that Odin-Sam1 and EphA2-Sam may adopt a mid loop/end helix binding model that is characteristic of Sam–Sam complexes.^{13,14,34,37}

SPR and ITC experiments were also performed, and both confirmed a clear association of Odin-Sam1 and EphA2-Sam (Figure 4). For SPR experiments, EphA2-Sam was efficiently immobilized on the chip surface while Odin-Sam1 was used as analyte; kinetic experiments along with a plot of *R*_{U,max} values of each experiment versus Odin-Sam1 concentration, both employing a 1:1 interaction model, provided a low micromolar dissociation constant value for the Odin-Sam1–EphA2-Sam complex (Figure 4A,B). Specifically, a *K*_D value of 5.5 ± 0.9 μ M could be obtained by best fitting of experimental data via nonlinear regression analysis (Figure 4B).

ITC experiments also indicated that Odin-Sam1 associated with EphA2 with a 1:1 stoichiometry and a *K*_D of 0.62 ± 0.04 μ M (Figure 4C).

NMR and SPR studies with EphA2-Sam mutants¹² were conducted to gain additional insight into the mode of binding of Odin-Sam1 to EphA2-Sam (Figures S1 and S2 of the Supporting Information). Details about the design of the two mutant proteins were previously described.¹² Briefly, in the triple (K38A/R78A/Y81S) EphA2-Sam mutant, we mutated residues located in helix $\alpha 5$ and the $\alpha 1$ – $\alpha 2$ loop that constitute part of the putative interaction surface of EphA2-Sam for Odin-Sam1 (i.e., regions undergoing major chemical shift changes upon binding of Odin-Sam1 to the receptor)

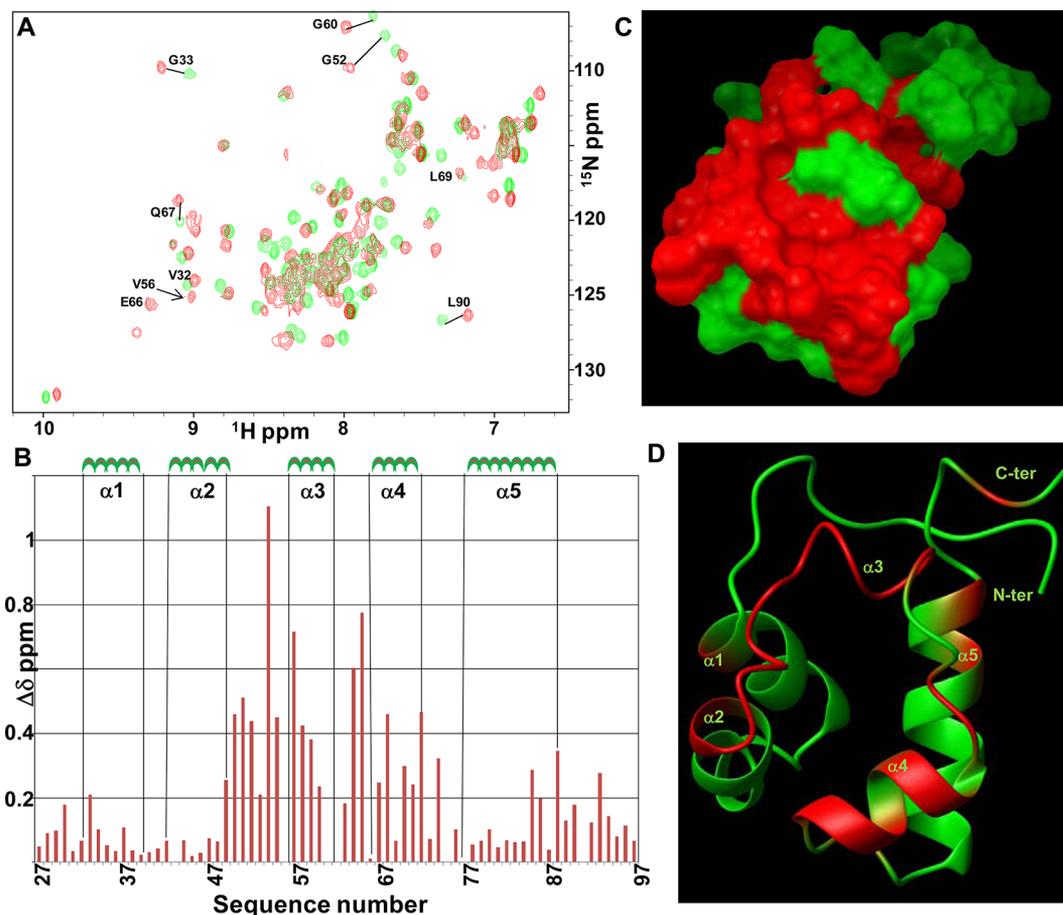


Figure 2. (A) Overlay of ^1H - ^{15}N HSQC spectra of Odin-Sam1 (150 μM) before (green) and after addition of EphA2-Sam (363 μM) (red). (B) Graph of chemical shift deviations $\{\Delta\delta = [(\Delta H_N)^2 + (0.17 \times \Delta^{15}N)^2]^{1/2}\}$ vs residue number. A $\Delta\delta$ value equal to 0 has been assigned to residues Q43, V56, and N62, whose peaks can be seen only in the spectrum of Odin-Sam1 bound to EphA2-Sam, S61 and S75 (unassigned), P77, and P91. (C and D) Surface (C) and ribbon (D) representations of Odin-Sam1 (conformer 1); residues with $\Delta\delta$ values higher than 0.2 ppm (i.e., G33, L49, L50, N51, G52, F53, D54, D55, H57, F58, L59, G60, M64, E65, Q67, D68, R70, D71, I72, I74, Q85, R88, and V93) are colored red.

(Figure 3C and Figure S1 of the Supporting Information), and amino acid replacements were planned to destroy potential key interactions at the dimer interface without perturbing the overall protein structure. In the H45N/R71A EphA2-Sam construct, mutations were instead inserted in regions adjacent to the putative binding site (Figure 3C and Figure S2 of the Supporting Information).

NMR and SPR experiments revealed for the triple (K38A/R78A/Y81S) mutant a binding affinity lower than that of the wild-type protein for Odin-Sam1 (Figure S1 of the Supporting Information). On the other hand, the EphA2-Sam double (H45N/R71A) mutant preserved an ability to associate with Odin-Sam1 similar to that of the wild-type protein (Figure S2 of the Supporting Information).

Molecular Modeling of the Odin-Sam1–EphA2-Sam Complex. A speculative model of the Odin-Sam1–EphA2-Sam complex was built by molecular docking with Haddock version 2.0^{32,38} (See Materials and Methods for details).

A representative structure (corresponding to the second best Haddock solution) is shown in Figure 5A. These modeling studies, in agreement with chemical shift perturbation data, indicated for the complex a head-to-tail topology also known as mid loop (ML)/end helix (EH) binding mode,^{13,14} in which Odin-Sam1 and EphA2-Sam are providing the ML and EH binding interfaces, respectively. Intermolecular interactions

mainly include a network of H-bonds and electrostatic contacts between positively charged residues of EphA2-Sam and negatively charged residues of Odin-Sam1 (Figure 5A and Figure S3 of the Supporting Information). Cation- π and π - π interactions also occur in a few Haddock solutions and involve mainly Phe58 on the surface of Odin-Sam1 and to a lesser extent Tyr81 (EphA2-Sam EH site) and Phe53 (Odin-Sam1 ML site) (Figure 5A and Figure S3 of the Supporting Information).

DISCUSSION

EphA2 receptor tyrosine kinase is considered a promising target in drug discovery for cancer therapies, and the process of receptor endocytosis has been exploited as a possible route to reduce tumor malignancy.¹ Recent evidence has shown that Sam domains are crucial for anchorage of protein regulators of endocytosis at the receptor site.^{8,9}

The Sam domain of lipid phosphatase Ship2 is engaged in a heterotypic interaction with the Sam domain of the receptor and is able to inhibit its endocytosis in cancer cells.⁸ Proteins of the ANKS family also play a prominent role in the process through their Sam domains, possibly by regulating ubiquitination mechanisms.⁹

In this study, we focus our attention on Odin,¹⁶ an ANKS family member that in cancer cells increases receptor stability.³⁹

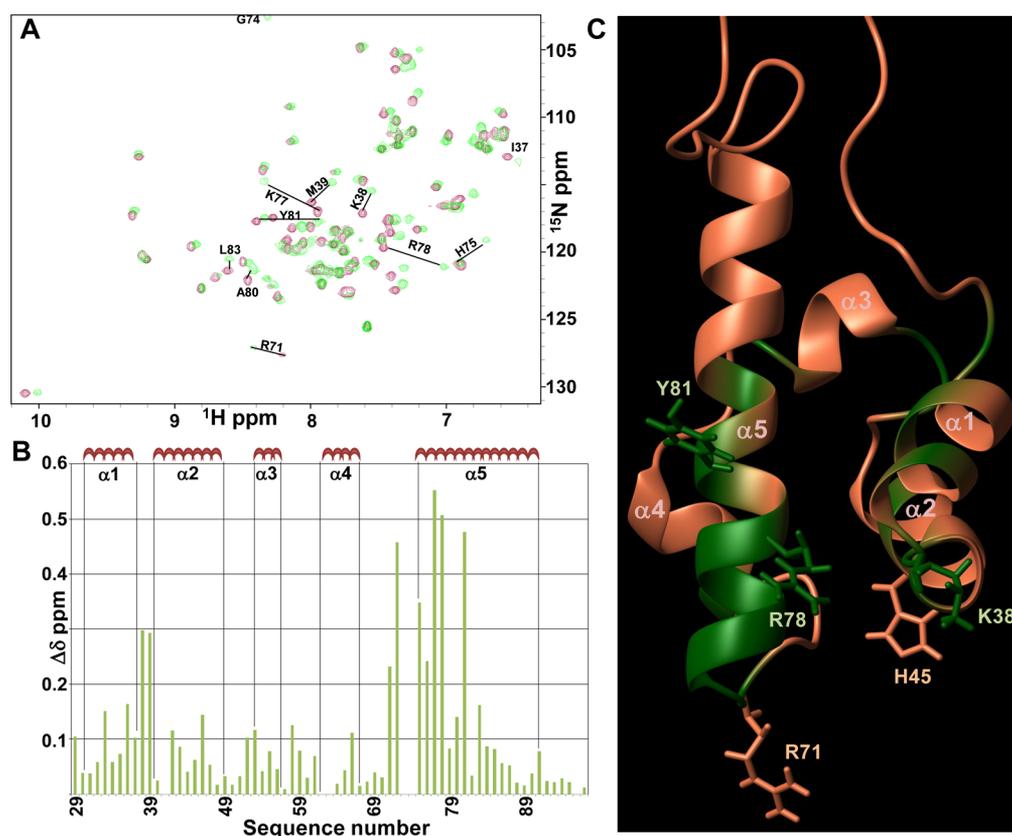


Figure 3. (A) Superposition of ^1H - ^{15}N HSQC spectra of EphA2-Sam ($63\ \mu\text{M}$) in the apo form (maroon) and bound to Odin-Sam1 ($330\ \mu\text{M}$) (green). (B) Histogram of normalized chemical shift deviations vs residue number. The largest variations ($\Delta\delta$ values >0.1 ppm) are observed for residues T29, W33, S36, I37, K38, M39, Y42, F46, T52, A53, V58, K66, R71, L72, H75, Q76, K77, R78, A80, Y81, and L83. Data are set equal to 0 for residues N62 and D63 (unassigned), G74 and Q41 (their peaks are visible only in the spectrum of the bound protein), P73, and P96. (C) Ribbon representation of EphA2-Sam (conformer 1, PDB entry 2E8N, from RIKEN Structural Genomics Initiative); residues with $\Delta\delta$ values >0.1 ppm are colored green. Side chains of amino acids involved in mutagenesis studies are colored green (EphA2-Sam triple mutant) and coral (EphA2-Sam double mutant).

Odin contains at the C-terminal side two Sam domains in tandem, Odin-Sam1 and Odin-Sam2. We attempted to express both isolated Sam domains (data not shown), but we could obtain only soluble Odin-Sam1 protein. However, difficulties in expression of the Sam2 domain of AIDA1b, another member of the ANKS protein family with a sequence highly homologous with that of Odin, were previously encountered.^{9,34} For Odin-Sam1, we conducted a complete structural characterization by NMR.

Sam domains exhibit a generally weak tendency to associate in solution through homotypic interactions.^{12,18,40} Our studies of the aggregation state of Odin-Sam1 confirm this trend. In fact, the correlation time of Odin-Sam1 ($\tau_c = 7.3$ ns at a protein concentration of $900\ \mu\text{M}$), evaluated by ^{15}N relaxation data, is rather close to that reported for monomeric Sam domains such as Ship2-Sam (6.7 ns)¹² and Arap3 (Arf GAP, Rho GAP, ankyrin repeat, and PH domain)-Sam (8.2 ns) at similar concentrations and under similar buffer conditions.¹⁸ The presence of one single Odin-Sam1 monomeric species in solution is further supported by DOSY experiments.²³ In particular, the hydrodynamic radius of the protein (i.e., ~ 17 Å) measured by DOSY²³ is comparable with that of compact proteins of a similar size.⁴¹

The correlation time of Odin-Sam1 bound to EphA2-Sam increases instead to ~ 11 ns, thus reflecting the increase in molecular weight upon association and suggesting that the two proteins may bind with a 1:1 stoichiometry;^{12,18,34} in fact, for

the AIDA1-Sam1/Sam2 tandem, a similar correlation time value equal to 9.1 ns has been reported.³⁴

To identify closely related Sam domains, we conducted a *blastp* search against the PDB⁴² by using the Odin-Sam1 sequence as the input query (Figure S4 of the Supporting Information). The results show that Odin-Sam1 presents the highest sequence identity (57%) with AIDA1-Sam1 (PDB entry 2EAM, RIKEN Structural Genomics Initiative); however, it possesses good homology also with Ship2-Sam (48% identical sequence, PDB entry 2K4P¹²) and EphA2-Sam (31% identical sequence, PDB entry 2E8N). A similar *blastp* search of Odin-Sam2 sequence indicates the highest identity with AIDA1-Sam2 (59%); sequence identities with EphA2-Sam and Ship2-Sam are instead 38 and 20% , respectively (Figure S4 of the Supporting Information).

In agreement with the high levels of homology revealed by *blastp*, Odin-Sam1 (Figure 1) presents the canonical Sam domain fold, and its structure is rather similar to that of AIDA1-Sam1 (PDB entries 2EAM and 2KE7); in fact, differences can be revealed only in the intrinsically disordered regions.

Interaction between Odin-Sam1 and EphA2-Sam and Comparison with Other Heterotypic Sam–Sam Associations. Because of the high sequence identity between Ship2-Sam and Odin-Sam1 as well as the common function of regulators of EphA2 endocytosis, we have investigated if Odin-Sam1 could directly bind EphA2-Sam as Ship2-Sam does.

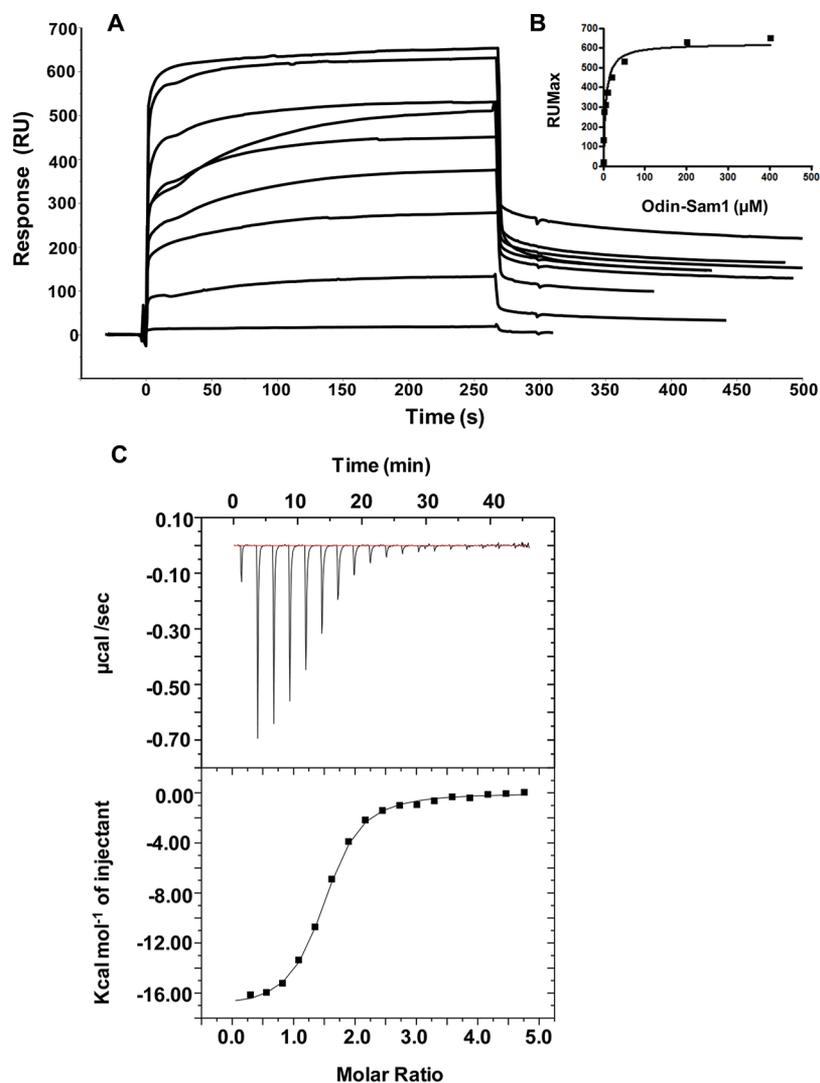


Figure 4. (A and B) SPR studies. Overlay of sensorgrams relative to the direct binding of Odin-Sam1 to immobilized EphA2-Sam (0.1–400 μM) (A). Plot of RU_{max} from each binding vs Odin-Sam1 concentration (B); data were fit by nonlinear regression analysis. (C) ITC studies. Calorimetric curve showing titration of EphA2-Sam (257 μM) with Odin-Sam1 (10 μM). The top and bottom sections report the raw and integrated data, respectively. A single-binding site model was applied for data fitting (bottom).

The interaction between Odin-Sam1 and EphA2-Sam has been studied by SPR, ITC, and NMR experiments. A clear association of the two proteins is evident from analysis of all the different binding assays. The dissociation constant estimated by SPR is in good agreement with that obtained by ITC (Figure 4). ITC data also clearly indicate, in agreement with the relaxation data, that the two Sam domains bind according to a single-binding site model (Figure 4C).

NMR chemical shift perturbation experiments (Figures 2 and 3) have revealed the reciprocal binding surfaces of the proteins and suggested that they may adopt a mid loop (ML)/end helix (EH) binding mode.⁴³ By having available the 3D structures of both Odin-Sam1 (PDB entry 2LMR) and EphA2-Sam (PDB entry 2E8N) and to recognize possible key intermolecular interactions stabilizing the complex, we have conducted molecular docking studies with the Haddock web server.³² Docking trials have been coupled to mutagenesis studies. The latter indicate that Tyr81, Lys38, and Arg78 on the EH surface of EphA2 are likely providing important interactions; in fact, concurrent mutations of Tyr to Ser and Lys and Arg to Ala attenuate the binding affinity for Odin-Sam1 (Figure S1 of the

Supporting Information). Indeed, in most of our docking solutions, these residues are involved in intermolecular contacts (Figure 5 and Figure S3 of the Supporting Information). The K38A, R78A, and Y81S mutations in EphA2-Sam also decrease the binding affinity of the receptor for Ship2-Sam.¹² We have previously reported the NMR solution structure of Ship2-Sam and characterized its binding to EphA2-Sam with NMR and ITC techniques.¹² These earlier studies have pointed out that Ship2-Sam and EphA2-Sam bind to each other by adopting a ML/EH binding topology;¹² very recently, the same interaction has been studied under different experimental conditions, and the structural details of the binding interface have been further elucidated.⁴⁴

The Ship2-Sam/EphA2-Sam binding mode and the type of possible stabilizing interactions^{12,44} appear similar to that we can observe for the Odin-Sam1–EphA2-Sam complex. Indeed, an NMR-based displacement experiment shows that Ship2-Sam and Odin-Sam1 share a common binding site on the surface of the receptor (Figure S5 of the Supporting Information).

Next, we compared our Odin-Sam1–EphA2-Sam model with the NMR structure of the tandem Sam domain of AIDA1 (PDB

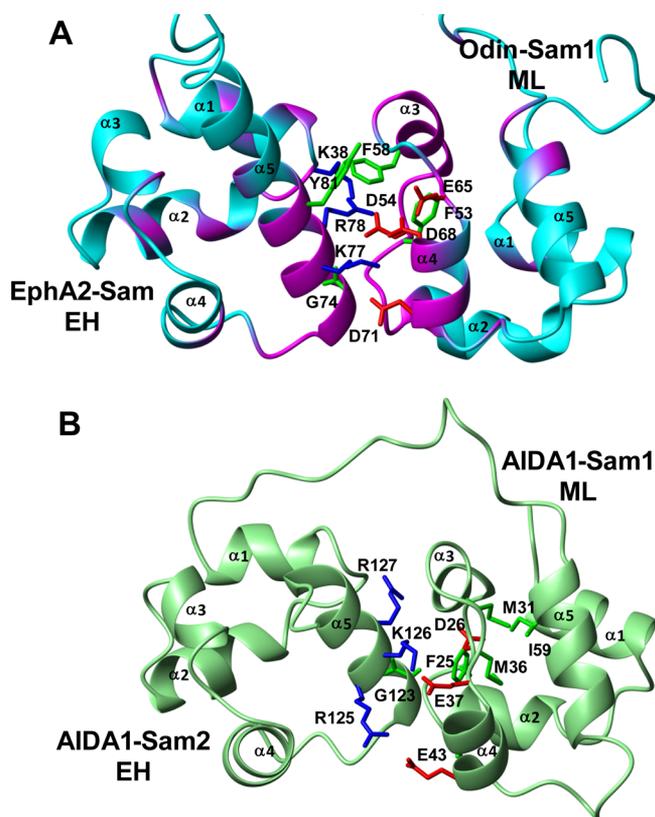


Figure 5. (A) Model of the Odin-Sam1–EphA2-Sam complex (second ranked Haddock structure belonging to the most populated docking cluster 1). Protein regions with the largest chemical shift variations upon Sam–Sam association, according to NMR chemical shift perturbation experiments, are colored magenta, and a subset of the amino acids providing interactions at the dimer interface are indicated. (B) NMR structure of the AIDA-1 Sam domain tandem (PDB entry 2KIV, structure n.1).³⁴ Most of the residues belonging to the ML and EH interfaces of Sam1 and Sam2 are shown.³⁴

entry 2KIV³⁴) (Figure 5B). Analogies between the two Sam–Sam associations are evident. Interestingly, the tandem presents a mid loop/end helix topology in which AIDA1-Sam1 and AIDA1-Sam2 are supplying the ML and EH binding surfaces, respectively³⁴ (Figure 5A).

On the basis of the high degree of sequence homology between AIDA1 and Odin (see above), we suppose that the two Odin Sam domains in tandem may bind with an analogue ML/EH topology in which Odin-Sam1 supplies the ML interface. It is clear that if the ML surface of Odin-Sam1 is engaged in the interaction with Odin-Sam2 in the full-length protein, uncoupling of the Sam2 domain from the tandem may be needed to permit binding of Sam1 to EphA2-Sam. Indeed, it has already been hypothesized that the opening of the AIDA1 tandem is required for translocation of AIDA1 toward the nucleus.³⁴

However, more structural and biochemical studies are needed to shed light on the complex mechanisms regulating the network of interactions of Odin and EphA2-Sam.

Moreover, it is worth noting that recent findings have associated a few mutations in EphA2-Sam with cataracts.^{6,45} Some of these EphA2-Sam mutations are located close to the EH interface (i.e., the $\alpha 4$ – $\alpha 5$ loop and the C-terminal portion of helix $\alpha 5$) (Figure S4B of the Supporting Information); whether they may cause perturbation in the structure of the

receptor Sam domain and thus influence its binding affinity for Odin-Sam1 remains to be addressed.

In this regard, it could be appealing to design, based on the structural information we have obtained here and in our previous work, novel molecular probes (small molecules or peptides) able to antagonize Odin-Sam1–EphA2-Sam association and study the outcomes in a cellular context, thus fully validating its relevance to either cancer or cataracts.

■ ASSOCIATED CONTENT

📄 Supporting Information

Details about chemical shift perturbation and SPR studies with EphA2-Sam mutants, figures showing several Haddock solutions and lists of intermolecular interactions, sequence alignments of different Sam domains, and figures related to the NMR displacement experiment with Ship2-Sam. This material is available free of charge via the Internet at <http://pubs.acs.org>.

Accession Codes

NMR structures of Odin-Sam1 have been deposited as Protein Data Bank entry 2LMR. Assigned chemical shifts of Odin-Sam1 have been deposited in the BioMagResBank as entry 18134.

■ AUTHOR INFORMATION

Corresponding Author

*Institute of Biostructures and Bioimaging, National Research Council, Via Mezzocannone 16, 80134 Naples, Italy. Phone: +39(081) 2534512. Fax: +39(081) 2536642. E-mail: marilisa.leone@cnr.it.

Funding

Financial support was obtained via National Institutes of Health Grant CA138390 to M.P. and FIRB Contract RBAP114AMK.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Mr. L. Zona, Mr. L. De Luca, and Dr. G. Perretta for technical assistance.

■ ABBREVIATIONS

Arap3, Arf GAP, Rho GAP, ankyrin repeat, and PH domain; AIDA1b, A β PP intracellular domain-associated protein 1b; AIDA1-Sam1, first Sam domain of AIDA1b; AIDA1-Sam2, second Sam domain of AIDA1b; ANKS, ankyrin repeat and sterile α motif domain protein; DOSY, diffusion-ordered spectroscopy; EH, end helix; EphA2, ephrin A2 receptor; EphA2-Sam, Sam domain of the EphA2 receptor; ITC, isothermal titration calorimetry; HSQC, heteronuclear single-quantum coherence spectroscopy; MD, mid loop; NOESY, nuclear Overhauser enhancement spectroscopy; Odin-Sam1, first Sam domain of Odin; Odin-Sam2, second Sam domain of Odin; PTB domain, phosphotyrosine binding domain; Ship2, Src homology 2 domain-containing phosphoinositide 5-phosphatase 2; Ship2-Sam, Sam domain of Ship2; SPR, surface plasmon resonance; TOCSY, total correlation spectroscopy.

■ REFERENCES

- (1) Pasquale, E. B. (2010) Eph receptors and ephrins in cancer: Bidirectional signalling and beyond. *Nat. Rev. Cancer* 10, 165–180.
- (2) Surawska, H., Ma, P. C., and Galgia, R. (2004) The role of ephrins and Eph receptors in cancer. *Cytokine Growth Factor Rev.* 15, 419–433.

- (3) Koolpe, M., Dail, M., and Pasquale, E. B. (2002) An ephrin mimetic peptide that selectively targets the EphA2 receptor. *J. Biol. Chem.* 277, 46974–46979.
- (4) Cooper, M. A., Son, A. I., Komlos, D., Sun, Y., Kleiman, N. J., and Zhou, R. (2008) Loss of ephrin-A5 function disrupts lens fiber cell packing and leads to cataract. *Proc. Natl. Acad. Sci. U.S.A.* 105, 16620–16625.
- (5) Jun, G., Guo, H., Klein, B. E., Klein, R., Wang, J. J., Mitchell, P., Miao, H., Lee, K. E., Joshi, T., Buck, M., Chugha, P., Bardenstein, D., Klein, A. P., Bailey-Wilson, J. E., Gong, X., Spector, T. D., Andrew, T., Hammond, C. J., Elston, R. C., Iyengar, S. K., and Wang, B. (2009) EPHA2 is associated with age-related cortical cataract in mice and humans. *PLoS Genet.* 5, e1000584.
- (6) Zhang, T., Hua, R., Xiao, W., Burdon, K. P., Bhattacharya, S. S., Craig, J. E., Shang, D., Zhao, X., Mackey, D. A., Moore, A. T., Luo, Y., Zhang, J., and Zhang, X. (2009) Mutations of the EPHA2 receptor tyrosine kinase gene cause autosomal dominant congenital cataract. *Hum. Mutat.* 30, E603–E611.
- (7) Lupberger, J., Zeisel, M. B., Xiao, F., Thumann, C., Fofana, I., Zona, L., Davis, C., Mee, C. J., Turek, M., Gorke, S., Royer, C., Fischer, B., Zahid, M. N., Lavillette, D., Fresquet, J., Cosset, F. L., Rothenberg, S. M., Pietschmann, T., Patel, A. H., Pessaux, P., Doffoel, M., Raffelsberger, W., Poch, O., McKeating, J. A., Brino, L., and Baumert, T. F. (2011) EGFR and EphA2 are host factors for hepatitis C virus entry and possible targets for antiviral therapy. *Nat. Med.* 17, 589–595.
- (8) Zhuang, G., Hunter, S., Hwang, Y., and Chen, J. (2007) Regulation of EphA2 receptor endocytosis by SHIP2 lipid phosphatase via phosphatidylinositol 3-kinase-dependent Rac1 activation. *J. Biol. Chem.* 282, 2683–2694.
- (9) Kim, J., Lee, H., Kim, Y., Yoo, S., Park, E., and Park, S. (2010) The SAM domains of Anks family proteins are critically involved in modulating the degradation of EphA receptors. *Mol. Cell. Biol.* 30, 1582–1592.
- (10) Kim, C. A., and Bowie, J. U. (2003) SAM domains: Uniform structure, diversity of function. *Trends Biochem. Sci.* 28, 625–628.
- (11) Meruelo, A. D., and Bowie, J. U. (2009) Identifying polymer-forming SAM domains. *Proteins* 74, 1–5.
- (12) Leone, M., Cellitti, J., and Pellicchia, M. (2008) NMR studies of a heterotypic Sam-Sam domain association: The interaction between the lipid phosphatase Ship2 and the EphA2 receptor. *Biochemistry* 47, 12721–12728.
- (13) Rajakulendran, T., Sahmi, M., Kurinov, I., Tyers, M., Therrien, M., and Sicheri, F. (2008) CNK and HYP form a discrete dimer by their SAM domains to mediate RAF kinase signaling. *Proc. Natl. Acad. Sci. U.S.A.* 105, 2836–2841.
- (14) Ramachander, R., and Bowie, J. U. (2004) SAM domains can utilize similar surfaces for the formation of polymers and closed oligomers. *J. Mol. Biol.* 342, 1353–1358.
- (15) Wei, Z., Zheng, S., Spangler, S. A., Yu, C., Hoogenraad, C. C., and Zhang, M. (2011) Liprin-mediated large signaling complex organization revealed by the liprin- α /CASK and liprin- α /liprin- β complex structures. *Mol. Cell* 43, 586–598.
- (16) Emaduddin, M., Edelmann, M. J., Kessler, B. M., and Feller, S. M. (2008) Odin (ANKS1A) is a Src family kinase target in colorectal cancer cells. *Cell Commun. Signaling* 6, 7.
- (17) Ghersi, E., Vito, P., Lopez, P., Abdallah, M., and D'Adamo, L. (2004) The intracellular localization of amyloid β protein precursor (A β PP) intracellular domain associated protein-1 (AIDA-1) is regulated by A β PP and alternative splicing. *J. Alzheimer's Dis.* 6, 67–78.
- (18) Leone, M., Cellitti, J., and Pellicchia, M. (2009) The Sam domain of the lipid phosphatase Ship2 adopts a common model to interact with Arap3-Sam and EphA2-Sam. *BMC Struct. Biol.* 9, 59.
- (19) Neri, D., Szyperki, T., Otting, G., Senn, H., and Wuthrich, K. (1989) Stereospecific nuclear magnetic resonance assignments of the methyl groups of valine and leucine in the DNA-binding domain of the 434 repressor by biosynthetically directed fractional ^{13}C labeling. *Biochemistry* 28, 7510–7516.
- (20) Grzesiek, S., and Bax, A. (1993) Amino acid type determination in the sequential assignment procedure of uniformly $^{13}\text{C}/^{15}\text{N}$ -enriched proteins. *J. Biomol. NMR* 3, 185–204.
- (21) Bartels, C., Xia, T., Billeter, M., Güntert, P., and Wüthrich, K. (1995) The program XEASY for computer-supported NMR spectral analysis of biological macromolecules. *J. Biomol. NMR* 6, 1–10.
- (22) Kay, L. E., Torchia, D. A., and Bax, A. (1989) Backbone dynamics of proteins as studied by ^{15}N inverse detected heteronuclear NMR spectroscopy: Application to staphylococcal nuclease. *Biochemistry* 28, 8972–8979.
- (23) Morris, K., and Johnson, C. S. (1992) Diffusion-ordered two-dimensional nuclear magnetic resonance spectroscopy. *J. Am. Chem. Soc.* 114, 3139–3141.
- (24) Price, W. S. (1997) Pulsed-field gradient nuclear magnetic resonance as a tool for studying translational diffusion: Part 1. Basic theory. *Concepts Magn. Reson.* 9, 299–336.
- (25) Talluri, S., and Wagner, G. (1996) An optimized 3D NOESY-HSQC. *J. Magn. Reson., Ser. B* 112, 200–205.
- (26) Kumar, A., Ernst, R. R., and Wuthrich, K. (1980) A two-dimensional nuclear Overhauser enhancement (2D NOE) experiment for the elucidation of complete proton-proton cross-relaxation networks in biological macromolecules. *Biochem. Biophys. Res. Commun.* 95, 1–6.
- (27) Herrmann, T., Güntert, P., and Wüthrich, K. (2002) Protein NMR structure determination with automated NOE assignment using the new software CANDID and the torsion angle dynamics algorithm DYANA. *J. Mol. Biol.* 319, 209–227.
- (28) Koradi, R., Billeter, M., and Wüthrich, K. (1996) MOLMOL: A program for display and analysis of macromolecular structures. *J. Mol. Graphics* 14, 29–32, 51–55.
- (29) Sanner, M. F., Olson, A. J., and Spehner, J. C. (1996) Reduced surface: An efficient way to compute molecular surfaces. *Biopolymers* 38, 305–320.
- (30) Johnsson, B., Lofas, S., and Lindquist, G. (1991) Immobilization of proteins to a carboxymethyl-dextran-modified gold surface for biospecific interaction analysis in surface plasmon resonance sensors. *Anal. Biochem.* 198, 268–277.
- (31) Rich, R. L., and Myszkowski, D. G. (2001) BIACORE J: A new platform for routine biomolecular interaction analysis. *J. Mol. Recognit.* 14, 223–228.
- (32) de Vries, S. J., van Dijk, M., and Bonvin, A. M. (2010) The HADDOCK web server for data-driven biomolecular docking. *Nat. Protoc.* 5, 883–897.
- (33) van Dijk, A. D., and Bonvin, A. M. (2006) Solvated docking: Introducing water into the modelling of biomolecular complexes. *Bioinformatics* 22, 2340–2347.
- (34) Kurabi, A., Brener, S., Mobli, M., Kwan, J. J., and Donaldson, L. W. (2009) A nuclear localization signal at the SAM-SAM domain interface of AIDA-1 suggests a requirement for domain uncoupling prior to nuclear import. *J. Mol. Biol.* 392, 1168–1177.
- (35) Pellicchia, M. (2005) Solution nuclear magnetic resonance spectroscopy techniques for probing intermolecular interactions. *Chem. Biol.* 12, 961–971.
- (36) Farmer, B. T. II, Constantine, K. L., Goldfarb, V., Friedrichs, M. S., Wittekind, M., Yanchunas, J. Jr., Robertson, J. G., and Mueller, L. (1996) Localizing the NADP $^{+}$ binding site on the MurB enzyme by NMR. *Nat. Struct. Biol.* 3, 995–997.
- (37) Stafford, R. L., Hinde, E., Knight, M. J., Pennella, M. A., Ear, J., Digman, M. A., Gratton, E., and Bowie, J. U. (2011) Tandem SAM Domain Structure of Human Casin1: A Presynaptic, Self-Assembling Scaffold for CASK. *Structure* 19, 1826–1836.
- (38) Dominguez, C., Boelens, R., and Bonvin, A. M. (2003) HADDOCK: A protein-protein docking approach based on biochemical or biophysical information. *J. Am. Chem. Soc.* 125, 1731–1737.
- (39) Zhuang, G., Brantley-Sieders, D. M., Vaught, D., Yu, J., Xie, L., Wells, S., Jackson, D., Muraoka-Cook, R., Arteaga, C., and Chen, J. (2010) Elevation of receptor tyrosine kinase EphA2 mediates resistance to trastuzumab therapy. *Cancer Res.* 70, 299–308.

(40) Thanos, C. D., Faham, S., Goodwill, K. E., Cascio, D., Phillips, M., and Bowie, J. U. (1999) Monomeric structure of the human EphB2 sterile α motif domain. *J. Biol. Chem.* 274, 37301–37306.

(41) Wilkins, D. K., Grimshaw, S. B., Receveur, V., Dobson, C. M., Jones, J. A., and Smith, L. J. (1999) Hydrodynamic radii of native and denatured proteins measured by pulse field gradient NMR techniques. *Biochemistry* 38, 16424–16431.

(42) Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res.* 25, 3389–3402.

(43) Kim, C. A., Gingery, M., Pilpa, R. M., and Bowie, J. U. (2002) The SAM domain of polyhomeotic forms a helical polymer. *Nat. Struct. Biol.* 9, 453–457.

(44) Lee, H. J., Hota, P. K., Chughra, P., Guo, H., Miao, H., Zhang, L., Kim, S. J., Stetzk, L., Wang, B. C., and Buck, M. (2012) NMR Structure of a Heterodimeric SAM:SAM Complex: Characterization and Manipulation of EphA2 Binding Reveal New Cellular Functions of SHIP2. *Structure* 20, 41–55.

(45) Shiels, A., Bennett, T. M., Knopf, H. L., Maraini, G., Li, A., Jiao, X., and Hejtmancik, J. F. (2008) The EPHA2 gene is associated with cataracts linked to chromosome 1p. *Mol. Vision* 14, 2042–2055.

(46) Laskowski, R. A., Rullmann, J. A., MacArthur, M. W., Kaptein, R., and Thornton, J. M. (1996) AQUA and PROCHECK-NMR: Programs for checking the quality of protein structures solved by NMR. *J. Biomol. NMR* 8, 477–486.

DOI: 10.1002/cbic.201200592

Heterotypic Sam–Sam Association between Odin–Sam1 and Arap3–Sam: Binding Affinity and Structural Insights

Flavia A. Mercurio,^[a] Daniela Marasco,^[a, b] Luciano Pirone,^[c] Pasqualina L. Scognamiglio,^[a] Emilia M. Pedone,^[b] Maurizio Pellecchia,^[d] and Marilisa Leone^{*[b]}

Arap3 is a phosphatidylinositol 3 kinase effector protein that plays a role as GTPase activator (GAP) for Arf6 and RhoA. Arap3 contains a sterile alpha motif (Sam) domain that has high sequence homology with the Sam domain of the EphA2-receptor (EphA2-Sam). Both Arap3-Sam and EphA2-Sam are able to associate with the Sam domain of the lipid phosphatase Ship2 (Ship2-Sam). Recently, we reported a novel interaction between the first Sam domain of Odin (Odin-Sam1), a protein belonging to the ANKS (ANKyrin repeat and Sam domain containing) family, and EphA2-Sam. In our latest work, we applied NMR spectroscopy, surface plasmon resonance (SPR) and

isothermal titration calorimetry (ITC) to characterize the association between Arap3-Sam and Odin-Sam1. We show that these two Sam domains interact with low micromolar affinity. Moreover, by means of molecular docking techniques, supported by NMR data, we demonstrate that Odin-Sam1 and Arap3-Sam might bind with a topology that is common to several Sam-Sam complexes. The revealed structural details form the basis for the design of potential peptide antagonists that could be used as chemical tools to investigate functional aspects related to heterotypic Arap3-Sam associations.

Introduction

Arap3 (Arf GAP with Rho GAP domain, Ankyrin repeat and PH domain 3) is a phosphatidylinositol 3 kinase (PI3K) effector protein that had originally been discovered by screening for novel binders of PtdIns(3,4,5)P3 (phosphatidylinositol (3,4,5)P3).^[1] Arap3 works as a GTPase activating protein for Arf6 and RhoA; it plays roles in biological processes associated with the formation of lamellipodia, cell adhesion and spreading, and regulation of actin cytoskeleton.^[2–5] In particular, recent reports have associated Arap3 with developmental angiogenesis^[6] and scirrhous gastric carcinoma.^[7]

The Arap3 sequence includes, among others, a sterile alpha motif (Sam) domain.^[8] Two proteins containing Sam domains, Ship2 (Src homology 2 domain-containing phosphoinositide-5-phosphatase 2) and ANKS1 (ANKyrin repeat and Sam-domain-containing 1), have been identified in a yeast two-hybrid screen as possible Arap3 regulators.^[4] The interaction between

Arap3 and Ship2 is mediated by Sam–Sam heterodimerization and has been well characterized.^[4,9] Arap3-Sam and Ship2-Sam bind to each other with a dissociation constant of approximately 100 nM;^[4] protein–protein association is mainly driven by specific electrostatic contacts.^[9] A docking model of the Ship2-Sam/Arap3-Sam complex, assisted by NMR and mutagenesis data, shows that the two proteins adopt a head-to-tail binding topology that is similar to that observed for many Sam–Sam associations; this is called the mid-loop/end-helix (ML/EH) model.^[4,10,11]

Similar analysis of the interactions between Arap3-Sam and Sam domains from the ANKS1 protein have not been reported so far. However, we have recently studied the heterotypic association between the first Sam domain of Odin (Odin-Sam1) and EphA2-Sam;^[12] this interaction is possibly relevant for EphA2 receptor endocytosis.^[13] Odin (also called ANKS1 and ANKS1A) is a protein belonging to the ANKS family, which has two tandem Sam domains (Sam1 and Sam2).^[14] In light of the relatively high sequence homology between EphA2-Sam and Arap3-Sam (~58%), we have investigated if Arap3-Sam could associate with Odin-Sam1. To verify and characterize binding between the two proteins we followed an approach comprising several techniques (NMR spectroscopy, isothermal titration calorimetry (ITC), surface plasmon resonance (SPR), molecular docking, and mutagenesis). We show that Arap3-Sam interacts with Odin-Sam1 with a dissociation constant in the low micromolar range and a 1:1 stoichiometry, and that it adopts a structural organization that closely resembles that of other ML/EH complexes, such as Ship2-Sam/Arap3-Sam^[9] and Odin-Sam1/EphA2-Sam.^[12]

[a] F. A. Mercurio, Dr. D. Marasco, Dr. P. L. Scognamiglio
Department of Biological Sciences, University of Naples "Federico II"
Via Mezzocannone 16, 80134 Naples (Italy)

[b] Dr. D. Marasco, Dr. E. M. Pedone, Dr. M. Leone
Institute of Biostructures and Bioimaging, National Research Council
Via Mezzocannone 16, 80134 Naples (Italy)
E-mail: marilisa.leone@cnr.it

[c] Dr. L. Pirone
Institute of Crystallography, National Research Council
Via Giovanni Amendola 122/O, 70126 Bari (Italy)

[d] Prof. M. Pellecchia
Infectious Diseases and Cancer Center
Sanford–Burnham Medical Research Institute
10901 North Torrey Pines Road, La Jolla, CA 92037 (USA)

 Supporting information for this article is available on the WWW under <http://dx.doi.org/10.1002/cbic.201200592>.

These studies provide novel and significant structural insights that should help the design of selective peptide/peptidomimetic molecules that could specifically antagonize Arap3-Sam heterotypic complexes. Such molecular probes could prove very useful in cell-based assays to shed light on the functional implications of these interactions.

Results and Discussion

Sam domains represent small helical protein-interaction modules of approximately 70 residues, and are characterized, in spite of a very similar fold, by high versatility concerning binding preference and function.^[8,15] Many Sam domain functions are mediated by the formation of homo- and heterotypic Sam-Sam interactions.^[16] From a structural point of view, Sam-Sam associations can adopt head-to-head, tail-to-tail, or even head-to-tail (ML/EH) topologies.^[9,17-21]

We previously characterized the associations between the Sam domains of the PI3K effector protein Arap3 (Arap3-Sam) and the lipid phosphatase Ship2 (Ship2-Sam),^[9] and more recently the interaction between the first Sam domain of Odin (Odin-Sam1) and the Sam domain of the EphA2 receptor (EphA2-Sam),^[12] and showed that these associations might adopt the head-to-tail topology. The current report focuses on characterization of the Odin-Sam1/Arap3-Sam interaction.

Odin-Sam1 binds Arap3-Sam: Chemical Shift perturbation Studies

To verify the association between Odin-Sam1 and Arap3-Sam we first conducted chemical shift perturbation studies with 2D ¹H,¹⁵N HSQC experiments.^[22] We acquired and compared NMR spectra of uniformly ¹⁵N-labeled Odin-Sam1, both in the free state and bound to unlabeled Arap3-Sam (Figure 1A). Association of the two Sam domains is clearly indicated by several changes in the HSQC spectra of Odin-Sam1 recorded in presence of increasing amounts of Arap3-Sam. Once saturation conditions had been reached (i.e., no change could be detected in the HSQC spectrum of Odin-Sam1 upon further addition of Arap3-Sam), the equation^[23] $\Delta\delta = [(\Delta H_N)^2 + (0.17 \cdot \Delta^{15}N)^2]^{1/2}$ was applied to evaluate proton and nitrogen normalized chemical shift deviations (Figure 1B) and to identify the interaction surface of Odin-Sam1 (Figure 2A). The greatest chemical shift perturbations ($\Delta\delta \geq 0.2$ ppm) were for the central portion of Odin-Sam1, encompassing helices α_3 and α_4 and the C-terminal portion of α_2 . Changes also occurred at the C-terminal tail, presumably because residues in this flexible portion come into proximity with the main binding interface (Figures 1B and 2A).

The pattern of chemical shift variations versus residue number observed for Odin-Sam1 in complex with Arap3-Sam

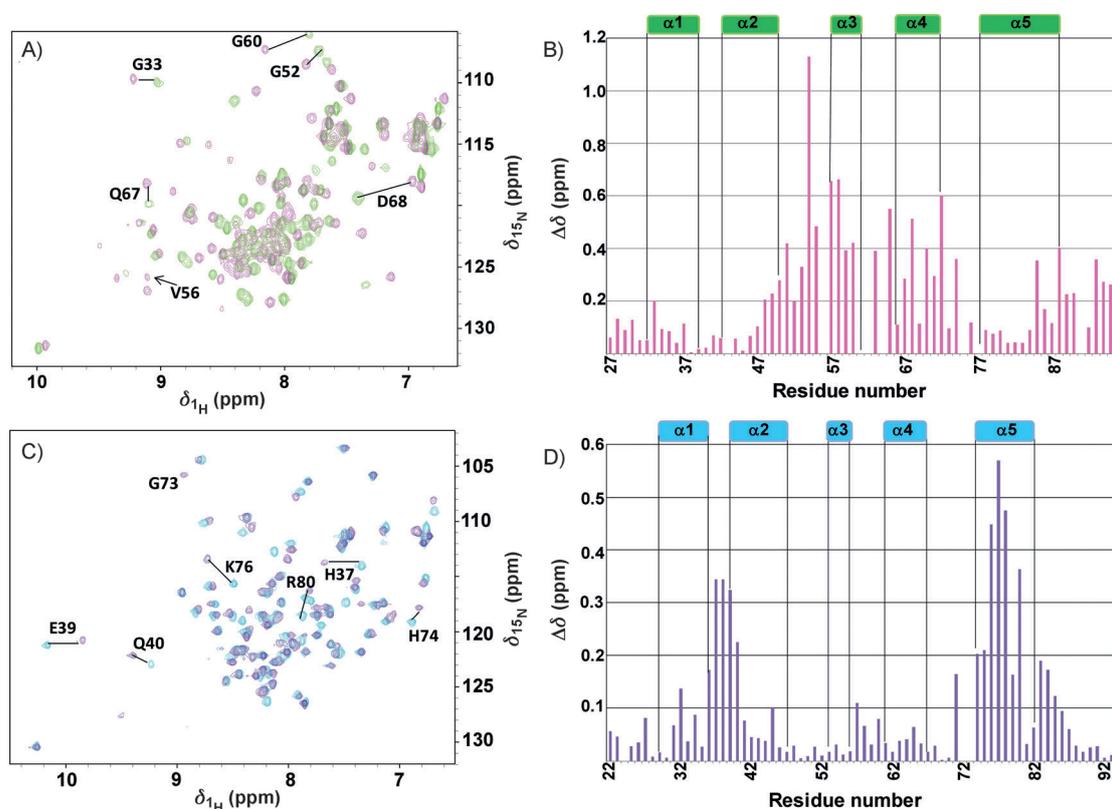


Figure 1. A) ¹H,¹⁵N HSQC spectra of ¹⁵N-labeled Odin-Sam1 (80 μM) in the free state (green) and in association with Arap3-Sam (370 μM; magenta). B) Chemical shift deviations ($\Delta\delta = [(\Delta H_N)^2 + (0.17 \cdot \Delta^{15}N)^2]^{1/2}$) against Odin-Sam1 sequence numbers. Residues Q43, V56, N62 (their peaks are revealed only in the spectrum of the Odin-Sam1/Arap3-Sam complex), S61, M64, S75 (unassigned), P77, and P91 were excluded from the analysis. C) Superposition of ¹H,¹⁵N HSQC spectra of ¹⁵N labeled Arap3-Sam (90 μM) in absence (cyan) and presence (violet) of unlabeled Odin-Sam1 (300 μM). D) Chemical shift deviations against Arap3-Sam residue numbers. Data were set equal to 0 for residues T72 and G73 (identified only in the spectrum of the bound protein), and P24.

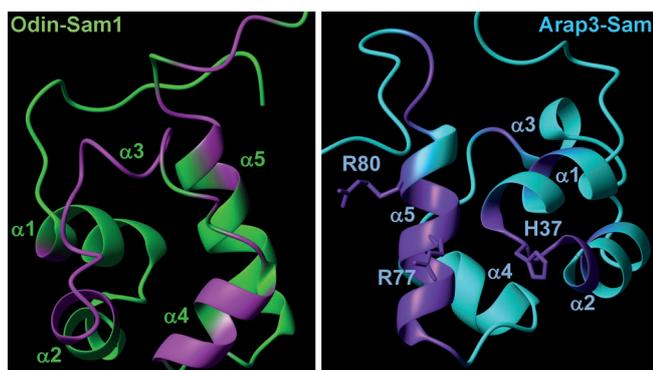


Figure 2. Left: The first conformer of Odin-Sam1 NMR ensemble (PDB ID: 2LMR)^[12] is shown in ribbon representation; backbone of residues with chemical shift deviations ≥ 0.2 ppm are highlighted in magenta (G33, L48, L49, L50, N51, G52, F53, D54, D55, V56, H57, F58, L59, G60, V63, E65, Q67, D68, R70, D71, I72, I74, Q85, R88, S89, L90, V93, K94, A95). Right: Ribbon representation of Arap3-Sam (NMR conformer number 1, PDB ID: 2KG5)^[9] residues with $\Delta\delta \geq 0.1$ ppm are colored purple (W32, V36, H37, L38, E39, Q40, F45, R57, A71, T72, G73, H74, R75, K76, R77, I78, L79, R80, Q83, T84, G85). Side chains of amino acids in the mutagenesis studies are shown.

(Figure 1B) closely resembles the one we previously detected for the protein in complex with EphA2-Sam.^[12]

NMR experiments with ¹⁵N-labeled Arap3-Sam and unlabeled Odin-Sam1 were also conducted, to map the binding surface of Arap3-Sam for Odin-Sam1 (Figure 1C and D). The largest chemical shift changes ($\Delta\delta \geq 0.1$ ppm) were for residues of helix $\alpha 5$ and the close $\alpha 1/\alpha 2$ and $\alpha 4/\alpha 5$ loop areas (Figures 1D and 2B). Interestingly, these are the regions that participate in the binding of Arap3-Sam to Ship2-Sam.^[9]

NMR perturbation data suggest that Odin-Sam1 and Arap3-Sam bind with a head-to-tail architecture, where the central region of Odin-Sam1 and helix $\alpha 5$ of Arap3-Sam, along with the adjacent loop regions, provide the main binding interfaces.

To get insights into the Odin-Sam1 and Arap3-Sam binding stoichiometry, we carried out ¹⁵N longitudinal (R1) and transverse (R2) nuclear spin relaxation rate measurements, evaluated the R2/R1 average values, and thus estimated the correlation time τ_c of both Sam domains bound to each other.^[24]

The τ_c of Arap3-Sam and Odin-Sam1 in their associated forms were (10 ± 1) and (9.9 ± 0.8) ns respectively. These values are of course higher than those for the proteins in their free states ((7.3 ± 0.6) ns for Odin-Sam1, (8.2 ± 0.4) ns for Arap3-Sam)^[9,12] because of the increased size of the complex and the consequent slower tumbling. Moreover, the τ_c estimates (ca. 10 ns) are also comparable to those evaluated for Odin-Sam1 in complex with EphA2-Sam,^[12] Ship2-Sam bound to Arap3-Sam (11 ns),^[9] as well as the AIDA1-Sam1/Sam2 tandem (9.1 ns),^[25] and point to a 1:1 binding stoichiometry for the Odin-Sam1/Arap3-Sam interaction.

SPR and ITC studies

To further quantify the Odin-Sam1/Arap3-Sam binding affinity, we performed SPR (Figure 3) and ITC experiments (Figure 4).

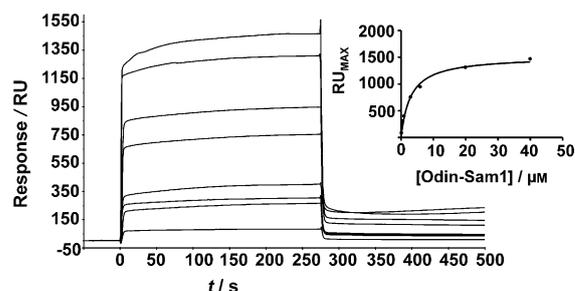


Figure 3. SPR Experiments. Superposition of sensorgrams corresponding to the interaction of Odin-Sam1 with immobilized Arap3-Sam (0.20–40 μM). Inset: plot of RU_{max} from each binding as function of Odin-Sam1 concentration; data were fitted by nonlinear regression analysis.

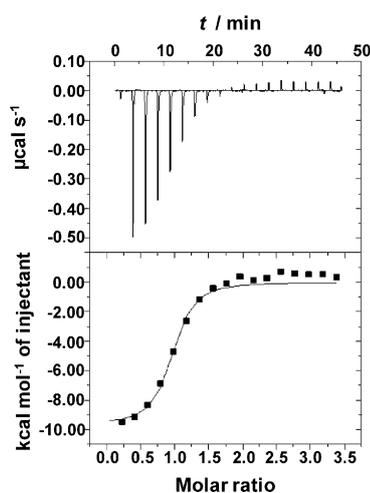


Figure 4. ITC Studies. Calorimetric curve for Odin-Sam1 (10 μM) titration with Arap3-Sam (250 μM). Raw and integrated data are shown in the upper and lower panels, respectively. In the lower section data fitting was achieved with a single binding site model.

SPR binding assays were carried out by immobilizing Arap3-Sam on the chip surface with Odin-Sam1 as the analyte. The fitting of data obtained by plotting RU_{max} values against Odin-Sam1 concentration, together with kinetic experiments implemented with a 1:1 binding model, gave a dissociation constant (K_D) in the low micromolar range (Figure 3). Indeed $K_D = (2.9 \pm 0.4)$ μM was obtained with a nonlinear regression analysis of experimental data (Figure 3).

In the ITC experiments (Figure 4), the exothermic heat peaks exhibited a monotonic decrease with addition of Arap3-Sam until saturation was reached. The data could be best fitted by a nonlinear least-squares approach to the “one set of sites” binding model (in agreement with ¹⁵N relaxation measurements, see previous paragraph), which yielded $K_D = (0.37 \pm 0.08)$ μM (Figure 4 and Table S1 in the Supporting Information). However, the estimated K_D values are similar to those determined in our previous binding assays for the Odin-Sam1/EphA2-Sam complex ((5.5 ± 0.9) μM and (0.62 ± 0.04) μM by SPR and ITC techniques, respectively),^[12] thus reflecting comparable binding affinities of Arap3-Sam and EphA2-Sam for Odin-Sam1.

**Structural features of the Odin-Sam1/Arap3-Sam complex:
HADDOCK models**

To shed light on the pattern of possible intermolecular contacts as the basis of the Odin-Sam1/Arap3-Sam interactions, we carried out molecular docking studies by using the HADDOCK server.^[26] These studies aimed at obtaining speculative models of the complex by using information from chemical shift perturbation studies and analysis of experimentally determined structures of other heterotypic Sam–Sam associations (Figure 5, and Section S2 in the Supporting Information). Our models present an ML/EH topology^[12,25] in which the mid-loop interface is provided by Odin-Sam1 and the end-helix surface is on Arap3-Sam (Figure 5C and D, Section S2).

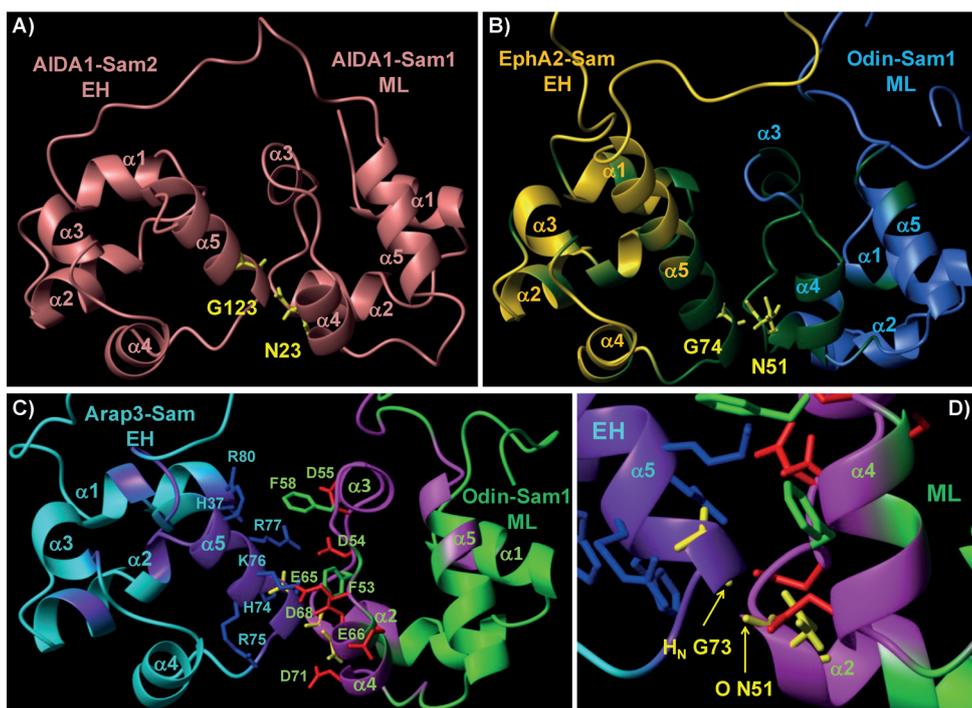


Figure 5. A) NMR structure of the AIDA-1-Sam domain tandem (PDB ID: 2KIV,^[25] structure 1). B) Representative HADDOCK^[26] model (number 2) of the Odin-Sam1/EphA2-Sam complex.^[12] The backbone of residues participating in the interaction, (according to NMR chemical shift perturbation data) are colored green.^[12] C) HADDOCK model (number 57) of Odin-Sam1/Arap3-Sam complex; residues that undergo major chemical shift variations following association are colored magenta and violet on the structures of Odin-Sam1 (green) and Arap3-Sam (cyan), respectively. Side chains of some amino acids that could contribute to the interaction interface are identified. D) Detail of the Odin-Sam1/Arap3-Sam HADDOCK model that better illustrates the H-bond between the backbone amide proton of Gly73 (Arap3-Sam EH site) and the backbone carbonyl oxygen of Asn51 (Odin-Sam1 ML binding region).

The Arap3-Sam binding region is rich in positively charged residues, whereas the Odin-Sam1 surface contains many negatively charged residues; thus, electrostatic interactions appear to be important for this association (Figure 5C and D).

To further investigate this, we studied the binding of Odin-Sam1 to an Arap3-Sam triple mutant (H37D, R77D, R80D).^[9] In this mutant we reversed the charge of three residues that are in the presumed Arap3-Sam EH binding site (Figure 2B). NMR and ITC assays failed to show significant binding of the triple

mutant to Odin-Sam1 (Figure S3) and led us to speculate that these residues supply pivotal interactions at the Sam–Sam interface and/or the mutations cause conformational changes in the binding region that inhibit association. Intriguingly, in most of the docking solutions, residue Arg77 in Arap3-Sam is engaged in electrostatic interactions with Asp54 and/or Asp55 in Odin-Sam1 (Figure 5C).

Arap3-Sam residues His37 and Arg80 seem less important, although in a few models His37 makes a salt bridge with Odin-Sam1 Asp55 (Section S2), while Arg80 could form a cation– π interaction with Odin-Sam1 Phe58 or a salt bridge with Odin-Sam1 Asp68 (Section S2). The Arap3-Sam mutant (H37D, R77D, R80D) was also unable to bind the Sam domain of the lipid phosphatase Ship2 (this can associate with wild-type Arap3-Sam by adopting an ML/EH binding model in which Arap3-Sam contributes the EH interface).^[9]

Another interaction to point out in the Odin-Sam1/Arap3-Sam docking models concerns residue Gly73, which is positioned on the EH surface of Arap3-Sam (Figure 5D and Section S2). It is interesting that the peak corresponding to the $^1\text{H}^{\text{N}}$ of Gly73 can only be detected in the $^1\text{H},^{15}\text{N}$ HSQC spectrum of Arap3-Sam in its bound form (Figure 1C), thus indicating that this Gly might become buried and possibly engaged in intermolecular contact in the Odin-Sam1/Arap3-Sam complex. Indeed, in a few docking solutions, the $^1\text{H}^{\text{N}}$ of Gly73 forms an intermolecular H-bond (Section S2). In particular, in the Odin-Sam1/Arap3-Sam model, the backbone H^{N} atom of Gly73 (Figure 5C and D) makes an H-bond with the backbone carbonyl oxygen atom of Odin-Sam1 residue Asn51. This kind of interaction between the backbone amide proton of a Gly (at the bottom of helix $\alpha 5$ on the EH interface) and the backbone carbonyl oxygen of another residue (at the C-terminus of helix $\alpha 2$ in an ML surface) is present in several experimentally determined structures of Sam–Sam complexes (see PDB IDs: 1PK1,^[27] 2KIV,^[25] 3BS5,^[20] 3SEI,^[10] 3SEN^[10]).

It has been proposed that a Gly residue at that position of an EH interface is highly desirable as it can facilitate the approach towards an ML binding region.^[25]

Indeed, recent NMR studies of the EphA2-Sam/Ship2-Sam complex revealed intermolecular NOEs between the backbone

H^N proton of Gly74 at the N terminus of helix $\alpha 5$ in the EphA2 receptor and the H β protons of an Asn at the C terminus of helix $\alpha 2$ in Ship2-Sam (PDB ID: 2KSO),^[28] thus further confirming our hypotheses.

Next, we compared our Odin-Sam1/Arap3-Sam model with the NMR structure of the AIDA1-Sam1/Sam2 tandem (PDB ID: 2KIV,^[25] Figure 5A) and found high homology. Like Odin,^[14] AIDA1^[29] belongs to the ANKS family; in the 3D structure of the AIDA1-Sam tandem, Sam1 and Sam2 bind by providing the ML and EH sites, respectively,^[25] and, interestingly, a Gly residue at the beginning of helix $\alpha 5$ on the binding interface of Sam2 likely provides an intermolecular H-bond with an Asn on the Sam1 interaction surface (Figure 5A). Moreover, our Odin-Sam1/Arap3-Sam models resemble the docking structures we recently proposed for the Odin-Sam1/EphA2-Sam complex (Figure 5B), in which EphA2-Sam provides the EH interaction surface, several intermolecular salt bridges can take place at the Sam-Sam interface, and a Gly at the N terminus of helix $\alpha 5$ (on the EH side) might be involved in an intermolecular H-bond.^[12]

In fact, a displacement assay with NMR techniques indicated that Odin-Sam1 employs the same ML interaction site to bind both EphA2-Sam and Arap3-Sam (Figure S4), and thus behaves like the Sam domain of the lipid phosphatase Ship2 (Ship2-Sam).^[9] The ML site of Ship2-Sam binds both Arap3-Sam and EphA2-Sam.^[9,19] The structural similarities between heterotypic associations of Arap3-Sam and EphA2-Sam probably reflect the high sequence homology between these two Sam domains (Figure S5) and appear to dictate a shared protein interactions network. As heterotypic Sam-Sam interactions of EphA2-Sam with either Odin and Ship2 are associated with EphA2 receptor endocytosis,^[13,30] we cannot exclude that Arap3-Sam might play a role in this process by sequestering two crucial regulators, but to date the intricate machinery that governs these heterotypic associations is not completely comprehended.

Conclusions

A yeast two-hybrid screen previously identified ANKS1 protein as a binding partner for the PI3K effector Arap3.^[4] The ANKS family includes Odin (also known as ANKS1 and ANKS1A) and AIDA-1 (ANKS1B), which have two Sam domains in tandem.^[14] The NMR structure of the tandem Sam domains of AIDA-1 reveals a characteristic Sam-Sam head-to-tail topology in which Sam1 and Sam2 bind with a mid-loop/end-helix model.^[10,25] The structure of Odin-Sam1 was solved by us with solution NMR techniques,^[12] whereas the experimental structure of the tandem Sam domains of Odin is not yet available in the Protein Data Bank. However, because of the high sequence similarity to AIDA-1, it is likely that the Odin Sam1-Sam2 tandem adopts a similar ML/EH architecture.

Herein, we investigated the interaction between Odin-Sam1 and Arap3-Sam and demonstrated that the two domains bind with a dissociation constant in the low micromolar range by forming a ML/EH heterodimer that is seemingly stabilized by electrostatic interactions. In our model, the central part of Odin-Sam1 and the C-terminal helix, together with close loop

regions of Arap3-Sam, supply the ML and EH binding sites, respectively. Taken together, our studies suggest that the interactions between Odin-Sam1 and Arap3-Sam might require separation of Odin-Sam2 from the Odin-Sam1-Sam2 tandem in the intact protein, similarly to our previous hypothesis for the association between EphA2-Sam and Odin-Sam1.^[12] Moreover, our observations, in conjunction with earlier data, suggest that Arap3-Sam adopts interaction patterns to Odin-Sam1 that highly resemble those between EphA2-Sam and Ship2-Sam. As both Odin and Ship2 are regulators of EphA2 receptor endocytosis,^[13,30] the revealed analogies might indicate an involvement of Arap3-Sam in this process. However, to date, the functional consequences of the association between Arap3-Sam and Odin-Sam1 are unknown. Thus, our immediate goals is to generate, based on our structural models, new peptide/peptidomimetic molecules that can interfere selectively with Arap3 Sam-Sam interactions, and analyse the resulting phenotypic alterations that such molecular probes induce in a cellular environment.

Experimental Section

Protein expression: Sam domains were expressed as recombinant proteins in *E. coli*. pET15B constructs encoding wild-type human Arap3-Sam (residues 1–80; UniprotKB/TrEMBL code: Q8WWN8),^[9] Arap3-Sam triple mutant (H37D, R77D, R80D),^[9] Odin-Sam1 (residues 691–770 of human Odin; UniprotKB/TrEMBL code: Q92625),^[12] and EphA2-Sam (residues 901–976 of human EphA2; UniprotKB/TrEMBL code: P29317)^[12] were purchased from Celtek Bioscience (Nashville, TN).

Genes were transformed into BL21-Gold (DE3) competent cells (Stratagene). Protein expression and purification procedures were conducted as previously described.^[9,12] In particular, to express unlabeled proteins, bacteria were grown in LB medium; ¹⁵N,¹³C doubly labeled and ¹⁵N-labeled proteins were expressed in M9 minimal medium containing [¹³C]glucose and/or [¹⁵N]NH₄Cl respectively. Bacteria were grown at 37 °C; β -D-thiogalactopyranoside (IPTG; 1 mM) was used to induce protein over-expression at a cell optical density OD₆₀₀ = 0.6 nm (overnight induction, T = 25 °C). Purification of His-tagged protein was performed on a nickel column with an ÄKTA Purifier apparatus (GE Healthcare).

Backbone resonance assignments of the Odin-Sam1/Arap3-Sam complex: NMR experiments were performed at 25 °C on a Varian Unity Inova 600 MHz spectrometer equipped with a cold probe.

HNCA and 3D ¹⁵N resolved ¹H,¹H NOESY spectra (100 ms mixing time), acquired with samples containing either ¹⁵N,¹³C doubly labeled Odin-Sam1 (620 μ M) and unlabeled Arap3-Sam (~3 mM), or doubly labeled Arap3-Sam (600 μ M) and unlabeled Odin-Sam1 (1.2 mM), were analyzed to obtain resonance assignments for the backbone H, N, and C α atoms of the Odin-Sam1/Arap3-Sam complex. NMR samples (600 μ L) were prepared in phosphate buffered saline (PBS; phosphate (10 mM), NaCl (138 mM), KCl (2.7 mM), pH 7.7; Fisher Scientific) with NaN₃ (0.2%) and D₂O (5%). Spectra were processed with VnmrJ software (version 1.1D; Varian) and analyzed with NEASY^[31] (included in the software package CARRA, <http://www.nmr.ch/>).

Relaxation measurements: NMR experiments to evaluate backbone ¹⁵N longitudinal (R1) and transverse (R2) relaxation rates, were performed at 25 °C on a Varian Unity Inova 600 MHz spec-

trometer provided with a cold probe. Measurements were carried out with two samples of Sam–Sam complexes: either ^{15}N , ^{13}C doubly labeled Odin-Sam1 (620 μM) plus unlabeled Arap3-Sam (3 mM), or ^{15}N , ^{13}C doubly labeled Arap3-Sam (100 μM) plus unlabeled Odin-Sam1 (~300 μM).

R1 and R2 relaxation data were collected as 1D spectra (4 K data points and 1–4 K transients). R1 data sets were recorded with the following relaxation delays: 0.01, 0.1, 0.3, 0.6, 1.0 s; R2 data set relaxation delays: 0.01, 0.03, 0.05, 0.07, 0.09, 0.11, 0.15, 0.19 s. Average R1 and R2 values were estimated by monitoring the decrease of signal intensity as function of the relaxation delays. The software package timest (A. G. Palmer III, Columbia University) was implemented to calculate the rotational correlation time from average R2/R1 ratios.^[32]

NMR binding studies: NMR titration experiments with 2D ^1H , ^{15}N HSQC spectra were carried out to study protein–protein interaction. The Odin-Sam1 binding interface for Arap3-Sam was identified by analysis of 2D ^1H , ^{15}N HSQC spectra of ^{15}N -labeled Odin-Sam1 (80 μM), in the unbound form and after addition of unlabeled Arap3-Sam (120, 370, and 620 μM). To recognize the binding site of Arap3-Sam for Odin-Sam1, 2D ^1H , ^{15}N HSQC spectra of a ^{15}N -labeled Arap3-Sam sample (90 μM) were recorded in the absence or presence of unlabeled Odin-Sam1 (100, 300, and 800 μM).

Binding of Arap3-Sam mutant (H37D, R77D, R80D) to Odin-Sam1 was investigated by analogous NMR chemical shift perturbation experiments conducted with either ^{15}N -labeled Arap3-Sam triple mutant (200 μM) and unlabeled Odin-Sam1 (200 and 300 μM), or ^{15}N -labeled Odin-Sam1 (144 μM) and unlabeled Arap3-Sam mutant (285 and 625 μM ; see the Supporting Information).

For the NMR displacement experiment, 2D ^1H , ^{15}N HSQC spectra were recorded for a ^{15}N -labeled Arap3 protein sample (50 μM) in its unbound form, after addition of Odin-Sam1 (150 μM), and in the simultaneous presence of Odin-Sam1 and EphA2-Sam (protein ratios after final dilutions: 1 (Arap3-Sam), 3 (Odin-Sam1), ~40 (EphA2-Sam)). Analysis of titration experiments and overlays of 2D spectra were generated with the program Sparky (version 3, T. D. Goddard and D. G. Kneller, University of California, San Francisco, <http://www.cgl.ucsf.edu/home/sparky/>).

Surface plasmon resonance: Arap3-Sam in acetate buffer (10 mM, pH 5.0) was immobilized (flow rate: 5 $\mu\text{L min}^{-1}$, injection time: 7 min) on a CM5 Biacore sensor chip by using EDC/NHS chemistry.^[33] Residual reactive groups were deactivated with ethanolamine hydrochloride (1 M, pH 8.5); the reference channel was prepared by activating with EDC/NHS and deactivating with ethanolamine. The immobilization level for Arap3-Sam was 1840 RU. Experiments were conducted at 25 °C and with a constant flow rate (20 $\mu\text{L min}^{-1}$), with a running buffer of HEPES (10 mM, pH 7.4), NaCl (150 mM), surfactant P20 (0.05% v/v); 90 μL injections for each experiment. Binding assays were conducted by using Odin-Sam1 at concentrations from 0.20 to 40 μM . The BIA evaluation analysis package (version 4.1, GE Healthcare, Milan, Italy) was used to subtract the reference channel signal and to evaluate K_D values. GraphPad Prism software (version 4.00; GraphPad Software, San Diego, California) was used to fit RU_{max} data against protein concentrations by nonlinear regression analysis.^[34]

ITC studies: ITC experiments were carried out with an iTC200 calorimeter (Microcal/GE Healthcare). Arap3-Sam (250 μM in PBS, pH 7.7) was titrated into a solution of Odin-Sam1 (10 μM in PBS, pH 7.7). Data were fitted to a single-binding-site model with Origin software (GE Healthcare). Similar ITC studies were conducted with

a solution of the Arap3-Sam triple mutant (H37D, R77D, R80D; 250 μM in PBS, pH 7.7) and Odin-Sam1 (10 μM in PBS, pH 7.7).

Molecular modeling: Docking studies were conducted with the HADDOCK web server.^[26] Models of the Odin-Sam1/Arap3-Sam complex were generated by starting from NMR structures (i.e., first conformer of Odin-Sam1: PDB ID: 2LMR,^[12] and first five structures of Arap3-Sam NMR ensemble: PDB ID: 2KG5).^[9] For Arap3-Sam, the flexible C-terminal tail (residues 90–100) was omitted from the docking procedure. Ambiguous interaction restraints were generated from chemical shift perturbation data. For both Odin-Sam1 and Arap3-Sam, active and passive residues were represented by a subset of those with higher normalized $\Delta\delta$, filtered by evaluating their solvent exposure and/or the possible involvement in intermolecular interactions, as suggested by analysis of experimental structures of Sam–Sam complexes (e.g., the AIDA-1 tandem Sam domain, PDB ID: 2KIV).^[25] For Odin-Sam1, residues L50, N51, F53, D54, D55, F58, E65, E66, D68, D71 were set as active, while L49, G52, V56, H57, S61, N62, V63, M64, Q67, R70 were considered passive; the region from L50 to E66 was set as “semiflexible” interface. Finally, the N- and C-terminal tails (M21 to G24 and A95 to N101, respectively) were set as “fully flexible” during all docking stages. For Arap3-Sam, the docking procedure included: G73, H74, K76, R77 as active residues; R75, L79 as passive residues; T72–R80 and V36–L38 as semiflexible tails; and H20–D26 as fully flexible region.

The solvated docking mode was implemented.^[35] One thousand structures were generated during the first phase of the docking protocol (rigid body energy minimization); next, semiflexible simulated annealing of the best 200 solutions was conducted; the last stage consisted of refinement in water. The final 200 HADDOCK solutions were visually analyzed, and “wrong” models (in disagreement with experimental NMR data, or presenting noncanonical Sam–Sam orientations) were discarded. After this first selection strategy, 106 HADDOCK models were further inspected with the MOLMOL software^[36] and compared with experimental structures of other heterotypic Sam–Sam complexes (reference PDB ID: 2KIV).^[25] to identify relevant structural features. From these, 23 models were considered representative of the possible conformations of the Arap3-Sam/Odin-Sam1 complex. Solutions were clustered by using a pairwise RMSD cutoff value of 2.26 Å and at least one structure per cluster. RMSD values were calculated with MOLMOL^[36] by superimposing the models on the backbone atoms of the secondary structure elements of both Arap3-Sam and Odin-Sam1 (i.e., residues 29–34, 39–47, 61–66, 72–83 for Arap3-Sam, and 32–39, 42–49, 57–60, 66–72, 77–88 for Odin-Sam1). Four clusters were obtained from the selected solutions (see Section S2 in the Supporting Information).

We also performed docking calculations with the more conventional HADDOCK protocol^[26] (Section S6). For these studies, we chose as active those exposed residues (solvent exposure evaluated with MOLMOL $\geq 24.6\%$) of Odin-Sam1 and Arap3-Sam with chemical shift variations ≥ 0.2 and 0.1 ppm, respectively. Passive residues were set automatically by HADDOCK; calculations were started from the initial ten conformers of both Sam domains; semiflexible and fully flexible regions were chosen as above. The results are discussed in the Supporting Information and are in good agreement with those obtained with the other described strategy, in which we used instead a reduced set of active and passive residues.

It is obvious that our models, built only on the basis of chemical shift perturbation data and analyzed by looking for structural analogies with other Sam–Sam heterocomplexes, need to be considered exploratory, as, of course, they do not have the precision of

a structure calculated by using unambiguous restraints such as intermolecular NOEs.

Acknowledgements

We thank Leopoldo Zona and Luca De Luca for technical assistance. Financial support was in part provided by NIH grant CA138390 to M.P. and FIRB Contract RBAP114AMK 006.

Keywords: isothermal titration calorimetry · NMR spectroscopy · protein–protein interactions · Sam domains · surface plasmon resonance

- [1] S. Krugmann, K. E. Anderson, S. H. Ridley, N. Risso, A. McGregor, J. Coadwell, K. Davidson, A. Eguinoa, C. D. Ellson, P. Lipp, M. Manifava, N. Ktistakis, G. Painter, J. W. Thuring, M. A. Cooper, Z.-Y. Lim, A. B. Holmes, S. K. Dove, R. H. Michell, A. Grewal, et al, *Mol. Cell* **2002**, *9*, 95–108.
- [2] S. Krugmann, S. Andrews, L. Stephens, P. T. Hawkins, *J. Cell Sci.* **2006**, *119*, 425–432.
- [3] T. T. I. Stacey, Z. Nie, A. Stewart, M. Najdovska, N. E. Hall, H. He, P. A. Ranzazzo, P. Lock, *J. Cell Sci.* **2004**, *117*, 6071–6084.
- [4] J. H. Raaijmakers, L. Deneubourg, H. Rehmann, J. de Koning, Z. Zhang, S. Krugmann, C. Erneux, J. L. Bos, *Cell. Signalling* **2007**, *19*, 1249–1257.
- [5] L. Gambardella, K. E. Anderson, C. Nussbaum, A. Segonds-Pichon, T. Margarido, L. Norton, T. Ludwig, M. Sperandio, P. T. Hawkins, L. Stephens, S. Vermeren, *Blood* **2011**, *118*, 1087–1098.
- [6] L. Gambardella, M. Hemberger, B. Hughes, E. Zudaire, S. Andrews, S. Vermeren, *Sci. Signaling* **2010**, *3*, ra76.
- [7] R. Yagi, M. Tanaka, K. Sasaki, R. Kamata, Y. Nakanishi, Y. Kanai, R. Sakai, *Oncogene* **2011**, *30*, 1413–1421.
- [8] C. A. Kim, J. U. Bowie, *Trends Biochem. Sci.* **2003**, *28*, 625–628.
- [9] M. Leone, J. Cellitti, M. Pellecchia, *BMC Struct. Biol.* **2009**, *9*, 59.
- [10] R. L. Stafford, E. Hinde, M. J. Knight, M. A. Pennella, J. Ear, M. A. Digman, E. Gratton, J. U. Bowie, *Structure* **2011**, *19*, 1826–1836.
- [11] R. Ramachander, J. U. Bowie, *J. Mol. Biol.* **2004**, *342*, 1353–1358.
- [12] F. A. Mercurio, D. Marasco, L. Pirone, E. M. Pedone, M. Pellecchia, M. Leone, *Biochemistry* **2012**, *51*, 2136–2145.
- [13] J. Kim, H. Lee, Y. Kim, S. Yoo, E. Park, S. Park, *Mol. Cell. Biol.* **2010**, *30*, 1582–1592.
- [14] M. Emaduddin, M. J. Edelman, B. M. Kessler, S. M. Feller, *Cell Commun. Signaling* **2008**, *6*, 7.
- [15] F. Qiao, J. U. Bowie, *Sci. STKE* **2005**, *2005*, re7.
- [16] A. D. Meruelo, J. U. Bowie, *Proteins Struct. Funct. Bioinf.* **2009**, *74*, 1–5.
- [17] C. D. Thanos, K. E. Goodwill, J. U. Bowie, *Science* **1999**, *283*, 833–836.
- [18] C. A. Kim, M. L. Phillips, W. Kim, M. Gingery, H. H. Tran, M. A. Robinson, S. Faham, J. U. Bowie, *EMBO J.* **2001**, *20*, 4173–4182.
- [19] M. Leone, J. Cellitti, M. Pellecchia, *Biochemistry* **2008**, *47*, 12721–12728.
- [20] T. Rajakulendran, M. Sahmi, I. Kurinov, M. Tyers, M. Therrien, F. Sicheri, *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 2836–2841.
- [21] M. Smalla, P. Schmieder, M. Kelly, A. Ter Laak, G. Krause, L. Ball, M. Wahl, P. Bork, H. Oschkinat, *Protein Sci.* **1999**, *8*, 1954–1961.
- [22] M. Pellecchia, *Chem. Biol.* **2005**, *12*, 961–971.
- [23] B. T. Farmer II, K. L. Constantine, V. Goldfarb, M. S. Friedrichs, M. Wittekind, J. Yanchunas, Jr., J. G. Robertson, L. Mueller, *Nat. Struct. Biol.* **1996**, *3*, 995–997.
- [24] N. A. Farrow, R. Muhandiram, A. U. Singer, S. M. Pascal, C. M. Kay, G. Gish, S. E. Shoelson, T. Pawson, J. D. Forman-Kay, L. E. Kay, *Biochemistry* **1994**, *33*, 5984–6003.
- [25] A. Kurabi, S. Brener, M. Mobli, J. J. Kwan, L. W. Donaldson, *J. Mol. Biol.* **2009**, *392*, 1168–1177.
- [26] S. J. de Vries, M. van Dijk, A. M. Bonvin, *Nat. Protoc.* **2010**, *5*, 883–897.
- [27] C. A. Kim, M. R. Sawaya, D. Cascio, W. Kim, J. U. Bowie, *J. Biol. Chem.* **2005**, *280*, 27769–27775.
- [28] H. J. Lee, P. K. Hota, P. Chugha, H. Guo, H. Miao, L. Zhang, S.-J. Kim, L. Stetzik, B.-C. Wang, M. Buck, *Structure* **2012**, *20*, 41–55.
- [29] E. Ghersi, P. Vito, P. Lopez, M. Abdallah, L. D'Adamo, *J. Alzheimer's Dis.* **2004**, *6*, 67–78.
- [30] G. Zhuang, S. Hunter, Y. Hwang, J. Chen, *J. Biol. Chem.* **2007**, *282*, 2683–2694.
- [31] C. Bartels, T.-h. Xia, M. Billeter, P. Güntert, K. Wüthrich, *J. Biomol. NMR* **1995**, *6*, 1–10.
- [32] L. E. Kay, D. A. Torchia, A. Bax, *Biochemistry* **1989**, *28*, 8972–8979.
- [33] B. Johansson, S. Löfås, G. Lindquist, *Anal. Biochem.* **1991**, *198*, 268–277.
- [34] R. L. Rich, D. G. Myszka, *J. Mol. Recognit.* **2001**, *14*, 223–228.
- [35] A. D. van Dijk, A. M. Bonvin, *Bioinformatics* **2006**, *22*, 2340–2347.
- [36] R. Koradi, M. Billeter, K. Wüthrich, *J. Mol. Graph.* **1996**, *14*, 51–55.

Received: September 13, 2012

Published online on December 13, 2012