AD_____

Award Number: W81XWH-05-1-0289

TITLE: Harnessing Novel Secreted Inhibitors of EGF Receptor Signaling for Breast Cancer Treatment

PRINCIPAL INVESTIGATOR: Mark A. Lemmon, Ph.D.

CONTRACTING ORGANIZATION: University of Pennsylvania
Philadelphia, Pennsylvania 19104-3246

REPORT DATE: April 2008

TYPE OF REPORT: Final

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

# REPORT DOCUMENTATION PAGE

*Form Approved*
*OMB No. 0704-0188*



| 1. REPORT DATE | 2. REPORT TYPE | 3. DATES COVERED |
|---|---|---|
| 01-04-2008 | Final | 7 MAR 2005 - 6 MAR 2008 |

**4. TITLE AND SUBTITLE**

Harnessing Novel Secreted Inhibitors of EGF Receptor Signaling for Breast Cancer Treatment

**5a. CONTRACT NUMBER**

**5b. GRANT NUMBER**
W81XWH-05-1-0289

**5c. PROGRAM ELEMENT NUMBER**

**6. AUTHOR(S)**
Mark A. Lemmon, Ph.D.

E-Mail: mlemmon@mail.med.upenn.edu

**5d. PROJECT NUMBER**

**5e. TASK NUMBER**

**5f. WORK UNIT NUMBER**

**7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)**

University of Pennsylvania
Philadelphia, Pennsylvania 19104-3246

**8. PERFORMING ORGANIZATION REPORT NUMBER**

**9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)**
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

**10. SPONSOR/MONITOR'S ACRONYM(S)**

**11. SPONSOR/MONITOR'S REPORT NUMBER(S)**

**12. DISTRIBUTION / AVAILABILITY STATEMENT**
Approved for Public Release; Distribution Unlimited

**13. SUPPLEMENTARY NOTES**

**14. ABSTRACT**

We aim to develop protein therapeutics that neutralizes growth factors that activate EGF receptor family members in breast cancer. Rather than targeting receptors themselves (as do Herceptin, Iressa, etc), we propose to target the activating ligands. Our model is Argos from Drosophila, which we showed naturally, inhibits EGF receptor signaling in fruit flies by inactivating the ligand. We hope to effectively 'humanize' Argos - making it bind human EGFR ligands and/or to use human protein scaffolds for this. In the past year, we crystallized a complex between the minimal functional fragment of Argos and its target (Spitz), and are about to complete structure determination – which will provide critical information for therapeutic design. We also established an experimental approach for screening libraries of Argos variants for those that bind human EGF-like ligands (our therapeutic aim). This approach employs yeast surface (rather than phage) display. We are now poised to combine our technical position and new structural information to identify Argos (and Dkk) variants that bind human EGFs and represent starting points for developing new therapeutics.

**15. SUBJECT TERMS**
EGF Receptor, Inhibitor, Antagonist, Argos, Ligand Sink, Signaling

| 16. SECURITY CLASSIFICATION OF: | | | 17. LIMITATION OF ABSTRACT | 18. NUMBER OF PAGES | 19a. NAME OF RESPONSIBLE PERSON |
|---|---|---|---|---|---|
| a. REPORT | b. ABSTRACT | c. THIS PAGE | | | 19b. TELEPHONE NUMBER *(include area code)* |
| U | U | U | UU | 64 | USAMRMC |

**Standard Form 298 (Rev. 8-98)**
**Prescribed by ANSI Std. Z39.18**

## Table of Contents

**Harnessing Novel Secreted Inhibitors of EGF Receptor Signaling for Breast Cancer Treatment**

## INTRODUCTION

The aim of this research project was to develop novel inhibitors of signaling through receptors from the epidermal growth factor (EGF) receptor family – well known to be important in breast cancer – by targeting (and neutralizing) the growth factor ligands rather than by targeting the receptors themselves. Apart from the relative success of Herceptin®, approaches that target ErbB receptors have yielded disappointing results in clinical trials, and this can be rationalized in retrospect. As described in a *Nature* letter [1] that we published in late 2004, we discovered that a physiologically important inhibitor of EGF receptor signaling in *Drosophila melanogaster* (named Argos) functions as a specific 'ligand-sink', neutralizing the fly's EGFR ligands and thus blocking signaling through this system. We proposed to develop a detailed understanding of how Argos achieves this 'ligand-sink' activity, and to analyze possible orthologs of these inhibitory molecules in humans and other organisms. Our ultimate goal was to modify the growth factor-neutralizing abilities of human orthologs of Argos (or structurally related proteins), and thus provide the essential groundwork for developing an innovative approach for breast cancer treatment that attacks the *ligands* responsible for activating EGF receptor family members rather than the receptors (which are the targets of all current therapies in this system – such as Iressa®, Tarceva®, lapatinib, Erbitux® and Herceptin®). One of the central problems is that each ligand activates multiple receptors in the family: through a combination of direct and indirect interactions. In cancers caused by aberrant ligand-induced ErbB receptor signaling, an effective therapeutic strategy would require targeting of all four receptors simultaneously (for which no agents exist), or – more straightforwardly – one would target the activating ligand(s). We proposed that our approach would lead to a novel (but physiologically validated – in *Drosophila*) approach for achieving this. As described in this report, although the precise route to achieving this goal is not quite what we predicted or proposed, we have identified human proteins that we hypothesize represent human Argos analogues [2]. Our next step is to investigate the possibility that these candidate human proteins can bind and sequester human ErbB receptor ligands.

Receptor tyrosine kinases (RTKs) from the EGF receptor family are well-validated therapeutic targets in breast and other cancers. The success of the ErbB2/HER2-targeted Herceptin® antibody [3] and more recently the ErbB2/EGFR tyrosine kinase inhibitor lapatinib [4] in breast cancer treatment has spurred efforts to achieve similar results with agents targeted to other receptors in this family. ErbB2/HER2/Neu is overexpressed to high levels in approximately 30% of breast cancer cases [5]. The EGF receptor is reported to be overexpressed in 14% - 91% of breast cancer patients [6]. It was therefore anticipated that therapeutic agents targeting the EGF receptor might show similar efficacy to Herceptin® in breast and other cancers [7]. However, although they looked very promising in preclinical studies [8, 9], EGF receptor inhibitors such as the cetuximab/Erbitux® antibody and tyrosine kinase inhibitors such as Iressa® and Tarceva® have yielded rather disappointing results in clinical trials [10-12], raising questions about whether the EGF receptor itself is really a good therapeutic target.

One of the key differences between ErbB2/HER2/Neu and the EGF receptor is that ErbB2 can be activated simply by over expression [13] (as occurs in breast cancer), whereas *EGF receptor still needs activating ligand even when over-expressed* [14]. It follows from this, and from the biology of ErbB receptors, that for clinical responses similar to those with Herceptin®, we should target the ligands of other receptors in the family rather than the receptors themselves. Developing a novel set of agents with this capability – based on our discovery that Drosophila use Argos to control their EGF receptor signaling in this way – is the aim of our current research.

## BODY OF FINAL REPORT

In our original application, we proposed three central strategies for identifying or generating Argos-like molecules that will function as 'ligand-sinks' for the many growth factors that activate human EGF receptor family members:

- analyzing the ability of distant Argos homologs in humans to bind human EGF-related ligands
- adaptation of human proteins that are structurally related to Argos for action as ligand sinks in signaling by the human EGF receptor family
- adaptation of *Drosophila* Argos to bind (and neutralize) human EGF-related ligands

Our conviction was that this combination of strategies, critically supported by structural work on Argos and the Argos/ligand complexes plus *in vitro* biophysical and cell biological studies, would provide starting points for developing drugs that can sequester the growth factors that activate EGF receptor family members in cancer (rather than targeting the receptors themselves). We believe that such agents will have significant advantages over other drugs that are currently in (disappointing) trials for EGF receptor inhibition in breast and other cancers.

### *Task 1.*
### To test the hypothesis that hDkks bind the EGF domains of EGF-related growth factors, and can act as inhibitory ligand 'sinks' for signaling by ErbB receptors

In the first year of the project, we made significant progress in this aim (with hDkk1 and hDkk3), and our results suggested that these molecules do not interact with human EGF-related ligands. As described in detail in the initial proposal, identification of the hDkks as human orthologs of Argos would have been the most straightforward (and ideal) route to development of molecules that can be used as therapeutic 'anti-EGFs'. In the initial application, we stated that:
*"It is quite possible (if not likely) that hDkk proteins will not bind to any of the human ErbB ligands, although the clear demonstration that hDkk1 and hDkk2 bind to EGF domains in LRP-6 makes this a reasonable hypothesis to test."*
The fact that our results showed that this simplest route would not be productive focused our attention more squarely on Aim 2 of the proposed research – to use hDkk proteins and Argos as 'scaffolds' for developing 'anti-EGF' proteins targeted against human ErbB ligands. Indeed, it appeared from sequence analyses that hDkk and Argos are structurally-related proteins that both bind EGF domains through their C-termini, making it seem a reasonable proposition that we might be able to alter their specificity. Progress towards this aim will be described below, although – as we published in *Nature* in May 2008 [2] – this approach was subsequently superceded by our finding that Argos has a highly unexpected structure that suggests a different set of potentially analogous proteins for future study.

Task 1a *Establish expression systems for the 4 human Dkk proteins in insect cells, and develop strategies for purification of near milligram quantities (months 1 to 4)*

We succeeded in expressing hDkk1 and hDkk3 by secretion from transfected *Drosophila* Schneider-2 (S2) cells as proposed (using the Invitrogen *Drosophila* Expression System). The proteins had their native signal sequences, and were tagged at their C-termini with a hexa-histidine tag. The proteins were secreted into the S2 cell medium, and were purified by passing dialyzed medium over a Ni-NTA column, followed by elution of bound protein with 100 mM imidazole, and subsequent gel filtration on a Superose-12 column. This purification procedure is identical to that used for Argos [1, 2]. The purified protein is homogeneous, as illustrated by the Western blot shown in Figure 1, and we succeeded in producing several hundred micrograms of material that appeared quite pure when analyzed by SDS-PAGE.

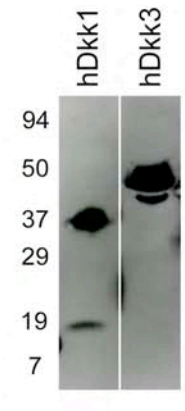


**Figure 1**
Western blot analysis of purified hDkk1 and hDkk3 purified from the conditioned medium of transfected S2 cells. Protein was detected using an antibody directed towards the C-terminal hexahistidine tag of the protein.
hDkk1 has a predicted mass of 27kDa (without carbohydrate), and contains one N-linked glycosylation site. hDkk3 has a predicted mass of 36kDa (without carbohydrate), and contains four potential N-linked glycosylation sites.

Task 1b *Using surface plasmon resonance, analyze binding of each hDkk protein to the 11 immobilized human ErbB ligands (months 4 to 6)*

We performed Biacore studies with these proteins using two 'formats'. In one, we generated biosensor surfaces on which hDkk1 or hDkk3 were immobilized. We could not detect binding of any human EGF-related growth factor to these surfaces. In the second approach, we generated biosensor surfaces on which any one of the 11 human EGF-related growth factors was immobilized, and passed solutions containing purified hDkk1 or hDkk3 across these surfaces to assess binding. Here, we have straightforward positive control experiments, since binding of the extracellular regions of EGFR or ErbB4 can be demonstrated for each ligand. In no case did we see convincing evidence for interaction of the hDkk proteins with human EGF-related growth factors, suggesting that the hypothesis tested in Aim 1 of the original proposal is incorrect.

Task 1c *Analyze the ability of each human Dkk protein to antagonize ErbB receptor activation by each ErbB ligand in a cellular context, and determine IC50 values for any positive cases (months 4 to 8)*

As anticipated given the outcome of Task 1b, in the proposed cellular studies, we saw no inhibition of growth factor-induced EGF receptor family member signaling upon addition of excess hDkk1 or hDkk3.

*Tasks 1a to 1c were all successfully completed* – although with negative results from the point of view of hDkk/ErbB ligand binding. Our negative findings in these sub-aims argue that our simplest route to identifying human Argos orthologs will not be productive. However, the likelihood of a positive result in these studies was remote at best, and these negative findings allowed us to focus our attention more squarely on Aims 1d and 2 of the proposed research – with structural studies of Argos (which have been very successful) and efforts to use hDkk proteins and Argos as potential 'scaffolds' for developing 'anti-EGF' proteins that can bind and neutralize human ErbB ligands.

Task 1d *Perform additional biophysical and structural characterization of any positive interactions between hDkk proteins and human ErbB ligands, using analytical ultracentrifugation and X-ray crystallography (months 8 to 24)*

We initially proposed using hDkk proteins and Argos as 'scaffolds' for developing novel antagonists for human ErbB ligands. To achieve this goal, a structural understanding of Argos itself, and its interactions with *Drosophila* ligands, is required. We have solved the crystal

structure of Argos alone and in complex with Spitz (its Drosophile EGF ligand), and the structures were all discussed in a paper that we published in *Nature* in May 2008 [2]. Although we anticipated that parallel studies of the structures of hDkk1 and hDkk3 would allow us to determine how these might best be used as scaffolds for Argos-like functions, it became clear from our success with crystallographic studies of Argos that the arguments for it sharing structural similarity with Dkks are not strong. As described below (and in our manuscript), Argos does not have an EGF domain as was first proposed. Moreover, the structure showed that Argos does not have a colipase fold after all – indicating that there is little reason to continue with our study of Dkk proteins. However, as discussed below (and in our *Nature* paper [2]), the Argos structure highlights the structurally-related (but poorly understood or characterized) C4.4A [15] and CD177/PRV-1 [16] proteins as potentially very exciting Argos analogues in humans.

Generating a form of Argos that behaved well enough to express and purify to high levels, and to crystallize, required significant protein engineering, guided largely by a combined genetic and biochemical analysis of Argos that we published in *J. Biol. Chem.* in 2006 [17]. These studies directed us to a 217-amino acid variant of Argos (the wild-type protein has 444 amino acids), which crystallized readily from 15% PEG 3400, 0.2M $(NH_4)_2SO_4$, 0.1 M KCl, 0.1M Acetate pH 4.6 (Figure 2). The crystals diffracted to 2.5Å resolution, and had space group C2 with a = 112Å, b = 62Å, c = 73Å and $\beta$ = 101°.

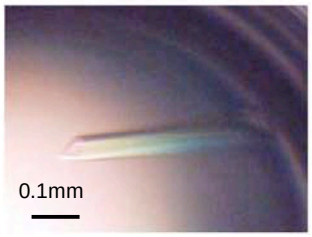


**Figure 2**
Example of a single crystal of $Argos_{217}$ grown as described in the text.

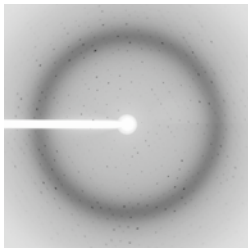

**Figure 3**
Representative image from a single $Argos_{217}$ crystal taken at APS beamline 23IDin, with 0.5° oscillation angle and a 3 second exposure. This crystal was grown in the presence of 7% glycerol, and frozen directly from the drop.

Solving the Argos structure from these crystals was very challenging. Although we had an excellent complete native dataset for the unliganded Argos crystals quite early on in the project, we were not able to obtain experimental phase information to solve the structure of this unliganded protein. Significant efforts to generate selenomethionine-containing protein failed, and the alternative approach of soaking in heavy atoms for multiple isomorphous replacement also failed to generate a useful heavy atom derivative, largely because stabilizing the Argos-alone crystals to allow reaction with mercury, platinum, etc compounds was not possible.

Dramatic progress with crystals of an Argos:ligand complex rescued this project, and led to some very exciting findings.

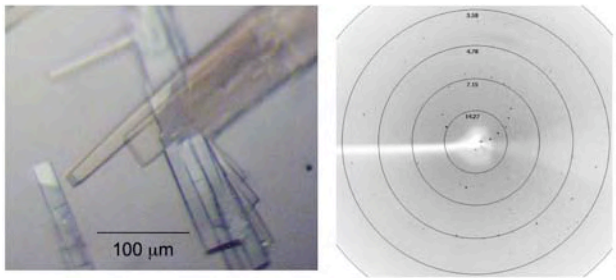


**Figure 4**

Crystals of an $Argos_{217}$/$Spitz_{EGF}$ complex (left), grown at 21°C from 4% PEG 20K, 100mM Hepes, pH 8.0. These crystals diffract to better than 2.5Å resolution on our home source, and to 1.6Å at synchrotron sources. The image at right shows diffraction of a $KPtCl_4$ derivative on our home source. See text.

Our best crystals of the $Argos_{217}$/$Spitz_{EGF}$ complex diffracted sufficiently well for us to have collected a complete dataset to 2.5Å resolution using our home X-ray source (Fig. 4). The crystals are of space group P1, with a = 50.4Å, b=52.1Å, and c=69.9Å, a=84.7˚, b=75.1˚ and g=77.1˚. These crystals are much more robust than the $Argos_{217}$-alone crystals, allowing us to soak them in a series of heavy atom solutions. We have obtained a clear platinum derivative of the $Argos_{217}$/$Spitz_{EGF}$ crystals, with one clear Pt site that refines well using both anomalous and isomorphous differences, in addition to several other potential sites. A preliminarily calculated SIR/AS map calculated using phases generated from this derivative looked quite promising – although additional phase information is required in order to generate a clearly interpretable map. During a synchrotron trip (at the Advanced Photon Source in Chicago), we measured diffraction to 1.6Å resolution, and succeeded in solving the structure by multiwavelength anomalous dispersion (MAD), using the halide soak method (Supplementary Table 1 of Klein et al. [2], attached in Appendix). Representative electron density is shown in Fig. 5a,b. Solving this complex structure allowed us to solve the structure of unliganded $Argos_{217}$ by sequential molecular replacement. In addition, the structure of uncomplexed $Spitz_{EGF}$ was solved to 1.5Å resolution by molecular replacement with a truncated human EGF model.

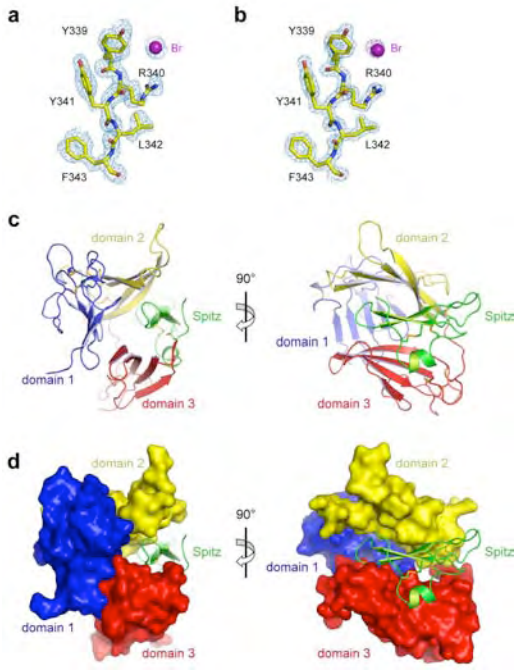


**Figure 5**

Structure of the Argos-Spitz complex.

**a**, Representative experimental electron density (contoured at 1σ) obtained after MAD phasing, showing a region of the Spitz-binding site on domain 2. The initial model is shown placed in the density. **b**, The same region of a 2Fo-Fc map (contoured at 2 σ) calculated using final phases. The final model is shown placed in the density. In purple, a peak corresponding to a bromide ion is seen in the anomalous difference Fourier map (contoured at 4 σ) using Br peak data. **c**, Cartoon of the $Argos_{217}$:$Spitz_{EGF}$ complex. Domains 1, 2 and 3 are colored blue, yellow and red respectively. Spitz is green. Disulphide bridges are colored orange. Two orthogonal views are shown. **d**, Same as in **c**, but with Argos in surface representation to emphasize the fact that it forms a clamp around the green Spitz molecule.

The structure showed that Argos consists of three separate disulphide-bonded β-sheet domains (domains 1-3) that are very closely related in structure (see below) and do not resemble EGF-like domains. This three-domain composition was not discerned by sequence analyses. As shown in Fig. 5c,d, the $Argos_{217}$/$Spitz_{EGF}$ complex has a trapezoidal shape, with one Argos domain at each of three corners and the bound ligand ($Spitz_{EGF}$) at the fourth. The average planes of the β-sheets formed by domains 2 and 3 of Argos are approximately parallel to one another, and both are nearly orthogonal to the plane of the domain 1 β -sheet. The three domains of Argos thus form a structure reminiscent of a C-clamp. Domains 2 and 3 constitute the 'jaws' of this clamp, and make an intimate set of direct contacts with bound $Spitz_{EGF}$. Domain 1 forms the backbone of the C-clamp and does not contribute directly to ligand binding. Domain 1 is separated from Spitz by a water-filled cavity.

Further details of the structure of Argos, the Argos/Spitz complex, and Spitz are discussed in our recently-published *Nature* letter [2], which is attached as an Appendix. Argos 'clamps' $Spitz_{EGF}$ between domains 2 and 3 (Fig. 5c,d), and buries 1360$Å^2$ of the ligand's

surface (35% of its total surface). Over half of the non-cystine side chains in $Spitz_{EGF}$ interact with Argos. Domains 2 and 3 are approximately parallel to one another, with domain 2 stacked above domain 3 in the complex (Fig. 5c), so they present opposite surfaces to the $Spitz_{EGF}$ molecule sandwiched between them. The three individual domains resemble hands with two fingers, and domain 3 presents its palm side to the bound Spitz molecule whereas domain 2 contacts Spitz using the back of the hand. Although it relies on a completely different structural scaffold, Argos mimics the characteristic bipartite capture of growth factor seen in ligand-bound structures of EGFR, and presents a ligand-binding surface that resembles the EGF-binding site in EGFR remarkably closely. Specifically, Argos domain 2 mimics domain I of sEGFR in its ligand contacts, whereas Argos domain 3 plays a similar role to sEGFR domain III in EGF domain binding (Fig. 6).

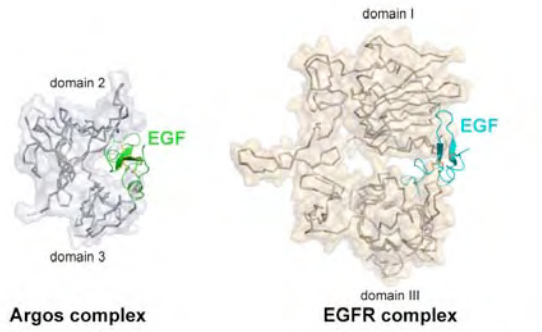


**Figure 6**

The left- and right- panels show EGF domains bound to Argos and the human EGFR extracellular region (sEGFR) respectively, illustrating the similarity in the 'bipartite capture' mode of EGF domain binding of both proteins. Spitz is coloured green and human EGF cyan.

The major advances represented by this structure determination come from the proteins that resemble Argos. These were completely unexpected. Sequence analyses have failed to identify clear mammalian homologues of Argos, but this does not necessarily mean that functional analogues do not exist in mammals. The amino acid sequence of Argos has been unusually cryptic, providing few (or misleading) clues about the structure of the protein. It only became apparent that Argos has three distinct domains (and no EGF-like domain) once the structure was determined. Moreover, the relationship of the constituent domains in Argos to the three-finger toxin fold can only been seen in structural (and not sequence) comparisons.

The individual domains of Argos share unexpected and striking structural similarity with the extracellular ligand-binding regions of receptors for TGFβ/bone morphogenetic protein (BMP) family ligands, which consist of little more than a single three-finger toxin fold similar to that seen in each of the three Argos domains [18, 19]. As shown in Fig. 7a (on next page), the two fingers of each Argos domain overlay well with two of the three fingers in TGFβ/BMP family receptor extracellular regions. The location of the disulphide-bonded core is also similar, although details of disulphide connectivity are different. The right-hand panel of Fig. 7a also shows how the positions of the ligand-binding sites in Argos domain 3 and the TGFβ/BMP family receptors correspond. Both utilize the palm side of the domain according to the hand analogy mentioned above.

Intriguingly, a trio of three-finger toxin fold domains engulf an EGF-like domain in another unexpected structural homologue of Argos: the cell-surface receptor for urokinase plasminogen activator (uPA) [20, 21]. The three domains from the uPA receptor (uPAR) form a clamp-like structure (Fig. 7b) around the EGF domain in the amino terminus of uPA, and each presents its palm side to the ligand. This structure resembles the clamp that is formed around the Spitz EGF domain by Argos (Fig. 5c,d). There are differences in the orientation of the bound EGF domain in the Argos/Spitz and uPAR/uPA complexes, and Argos has one of its three constituent domains (domain 2) 'inverted' so that it presents the back (rather than palm) of the hand to the ligand. However, the correspondence in overall architecture and function (as proteins that entrap EGF domains) of uPAR and Argos suggest that other structural homologues of Argos should be sought in mammals. There are many human uPAR/Ly6 domain-containing proteins for which the function remains unclear. Several, such as CD177/PRV-1 and C4.4A, contain multiple three-finger domains [15, 16] like uPAR. Moreover, C4.4A expression is known to be

altered in several metastatic human cancers [22]. We suggest that one of these numerous structural homologues might represent a functional analogue of Argos. Even if such an analogue does not exist, the known human proteins from this class could clearly be used as structural scaffolds in the design of protein therapeutics that will sequester ErbB receptor-activating EGF domains.

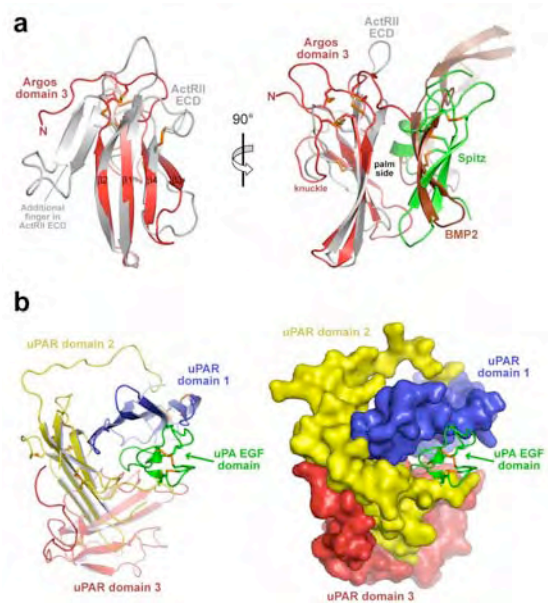


**Figure 7**
Similarity of Argos to TGFβ-family and uPA receptors **a**, Domain 3 of Argos (red) is overlaid with the 100aa extracellular ligand-binding domain of the type II activin receptor (ActRII)[23] (light gray). The two fingers of each Argos domain overlay well with the longest fingers of the ActRII three-finger toxin fold. In addition, the disulphide-bonded cores are similarly located. The right-hand panel shows an orthogonal view in which Spitz and the receptor-proximal region of BMP2 (bound to ActRII) are shown. Both ligands bind to the palm side of their respective binding domains. The 'third' finger of the ActRII ECD has been removed from this view for clarity. The opposite end of the extended BMP2 molecule binds to a similar site on a type I receptor. **b**, Domain organization of the uPA receptor. The three three-finger toxin fold domains in uPAR are colored with the same scheme used for Argos, and 'clamp' the green EGF domain of uPA that binds a central cavity. The molecule is shown in cartoon representation (left) and surface representation (right). Like Argos, uPAR uses three copies of this domain type to form a C-clamp-like structure for enveloping an EGF domain.

This structural analysis focuses our attention on determining whether C4.4A and CD177/PRV-1, or other proteins with multiple 3-finger toxin domains, might actually function like Argos in humans. We are now in the process of cloning these out in order to test this intriguing hypothesis. With a human Argos-like molecule in hand, the prospects for a new therapeutic agent are clear. If none bind human EGFR ligands, then we will have succeeded in identifying an excellent structural homologue that can benefit from the approached we have developed in Task 2 for development of human anti-EGF proteins.

**Task 2.**
**To adapt human Dkk proteins and *Drosophila* Argos to bind and neutralize human ErbB ligands**

In parallel with our structure determination studies, we focused our attention on engineering or evolving Argos to interact with the human EGF-related growth factor. In the original proposal, we described approaches for displaying the basic Argos and Dkk scaffolds as pIII fusions on M13 phage, so that we could select from limited randomized libraries to identify forms of Argos and hDkk that bind human EGF (and other ErbB ligands) with high affinity. As with all phage display projects, one of the first challenges is to display the protein on the phage surface while maintaining function. Since this is usually best achieved with the smallest fragment (or domain) of the protein that retains function, we completed Task 2b first – establishing that $Argos_{217}$ is the minimal Argos fragment that maintains functionality. Simultaneously, our structural studies made it clear that Argos is not likely to be structurally related to the hDkks after all, so we proceeded with $Argos_{217}$ – with a view that these experiments are 'proof-of-principle' experiments for uPAR, C4.4A and CD177/PRV-1 should this prove necessary (*i.e.* if none of these are actually human Argos analogs).

In the first year of the project, we put a great deal of effort into displaying Argos and other relevant proteins as pIII fusions on M13 phage (Task 2a). This was technically very challenging, possibly because the protein does not fold correctly in this context. This difficulty has resulted in a set-back, but one that we subsequently overcame by utilizing instead a yeast surface display approach pioneered by Wittrup and colleagues [24]. As described below, we succeeded in displaying Argos, Spitz, and human EGF on the surface of yeast cells. This provides the 'proof-of-principle' that this approach can be used for screening approaches. Rather than proceeding with a screen involving Argos (which – as a fly protein – will never bring us directly to a protein therapeutic with great utility in humans) we held off until deliberation with our Argos structural data suggested good *human* candidate proteins with which to start. We now have these in uPAR, C4.4A and CD177/PRV-1, and plan to initiate yeast display approaches as soon as possible after determining whether any of these binds human ErbB ligands without modification.

Task 2a*i* *Establish procedures for displaying the C-terminal cysteine-rich regions of hDkk1, hDkk2 and Argos as pIII fusions on M13 phage, and ensure that they are correctly presented by using phage ELISA assays to check binding to positive controls (LRP6 for hDkk1 and hDkk2; Spitz for Argos) (months 1 to 4)*

In order to maximize success of this task, and to display the most 'well-behaved' protein species on phage, we chose to complete Task 2b before Task 2a. As described below, we showed that an Argos fragment ($Argos_{217}$) could maintain full functionality, but with minimal glycosylation, suggesting that it should serve as an ideal target for yeast or phage display. We were not able to display $Argos_{217}$ on phage, although yeast approaches met with more success (see below).

Task 2b *Generate a set of nested deletions to determine what are the smallest C-terminal cysteine-rich fragments of hDkk1/2 and Argos that will bind to the positive control targets, and test binding by phage ELISA (months 4 to 8)*

Prior to our structure determination, by comparing the *D. melanogaster* Argos sequence with those from *M. domestica* and *A. mellifera*, we chose – after much trial-and-error – to delete residues 22-111 and 165-280 from the pre-protein (1-21 constitute the signal sequence). These deletions remove the blue parts of the primary structure as schematized in Fig. 8, which are largely unconserved regions that have little predicted secondary structure and contain multiple potential O-glycosylation sites.


**Figure 8**
Schematic of Argos primary structure. Mutants identified in Task 2d are noted.

This deletion mutant, with the green regions of Fig. 8 fused directly to one another to yield a 217aa protein (which we term $Argos_{217}$) expresses well, has excellent chromatographic properties, and the purified protein appeared homogeneous by SDS-PAGE (Fig. 9). $Argos_{217}$ and binds strongly to Spitz and the Spitz EGF-like domain. $K_D$ for Spitz binding is ~10nM (Fig. 9), which is very similar (in fact perhaps slightly stronger) to what we have previously published for Spitz binding by full-length Argos [1]. Thus, $Argos_{217}$ appears to have the same functional properties as the full-length secreted protein, but is much more robust – as required for phage display. Further deletion from $Argos_{217}$ impairs function. The structural studies described above were all completed with $Argos_{217}$, and the structure provides a clear rationale for why these particular deletions worked. We were unfortunately not able to display $Argos_{217}$ on phage, despite the fact that the protein is more robust and stable than the full-length wild-type protein.

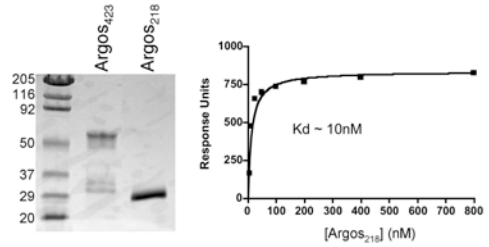


**Figure 9**
In the left-hand panel, a Coomassie-stained SDS-PAGE gel of full-length $Argos_{423}$ and the $Argos_{217}$ deletion mutant is shown, illustrating the dramatic improvement in homogeneity in the deletion mutant.
In the right-hand panel, binding of $Argos_{217}$ to the immobilized Spitz EGF-like domain is shown as a representative saturation-binding curve. The Spitz-binding properties have not been affected by the deletions.

These *in vitro* studies clearly identified the smallest cysteine-rich fragment of Argos that will bind to the positive control target (Spitz) – using purified protein rather that phage display as initially proposed.

Task 2a*ii* *Establish procedures for displaying the C-terminal cysteine-rich regions of hDkk1, hDkk2 and Argos as pIII fusions on M13 phage, and ensure that they are correctly presented by using phage ELISA assays to check binding to positive controls (LRP6 for hDkk1 and hDkk2; Spitz for Argos) (months 1 to 4)*

As mentioned above in the face of difficulties displaying Argos-derived proteins on phage, we turned to the yeast-surface display approach developed by Wittrupp and colleagues [24, 25], which is sold commercially by Invitrogen (as the pYD1 Yeast Display Vector Kit). Proteins of interest are expressed on the surface of *S. cerevisiae*, fused to the Aga2p protein (a component of the **a**-agglutinin receptor). The Aga2p portion of the fusion protein associates through two disulfide bonds with the Aga1p subunit of the **a**-agglutinin receptor, which is itself covalently attached to the outer surface of the yeast cell wall. As a result, the 'test' protein fused to Aga2p (in this case our Argos or hDkk) is covalently associated with the yeast surface. This approach has been validated in studies to evolve regions of the EGFR extracellular region to bind more strongly to anti-EGFR antibodies [26]. Moreover, Jin *et al*. have recently managed to gain a 200,000-fold increase in affinity of an integrin to ICAM-1 with this method [27]. The idea in this case is to perform selection (using the same randomized libraries outlined in the original proposal) using flow cytometry to select yeast clones that are specifically bound by a fluorescently labeled form of the target (human EGF, for example).

As shown in Figure 10, we can readily display either human EGF or the Drosophila Spitz ligand on the yeast surface using this approach. The surface-displayed EGF or Spitz is readily detected using flow cytometry (Fig. 10A) or immunofluorescence (Fig. 10B) with the relevant anti-EGF (or anti-Spitz) antibodies.

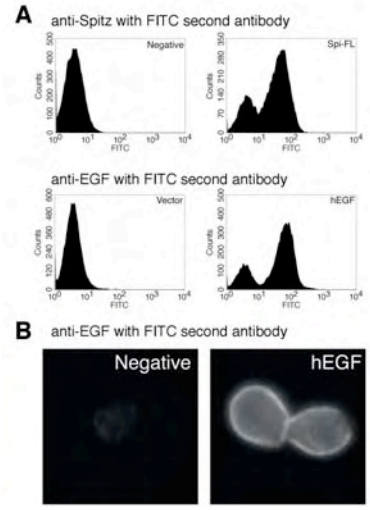


**Figure 10**
**A.** Flow cytometry analysis of EBY100 *S. cerevisiae* with anti-Spitz antibody (upper panels) or anti-human EGF antibody (lower panels) detected with FITC-labeled anti-IgG. Cells transfected with derivatives of plasmid pYD1 directing expression of a Spitz/Aga2p chimera (upper right panel) show strong staining with the anti-Spitz antibody. Cells transfected with the EGF-containing pYD1 derivative show strong staining with anti-EGF antibody.

**B.** The EGF-Aga2p fusion is detected clearly on the surface of yeast cells by immunofluorescence with anti-EGF antibodies.

These data demonstrate that we can readily express EGF-like ligands on the surface of *S. cerevisiae*. EGF and Spitz can also be detected at the cell surface using fluorescently-labeled versions of the *Drosophila* and human EGFR extracellular regions.

We have also shown that the human EGF displayed on the yeast surface is active in binding to its receptor, as is Spitz. Fig. 11 shows that yeast expressing hEGF fused to Aga2p bind both to a FITC-conjugated antibody against the histidine tag that it includes and to AlexaFluor-633 labeled sEGFR (the extracellular region of the human EGFR).

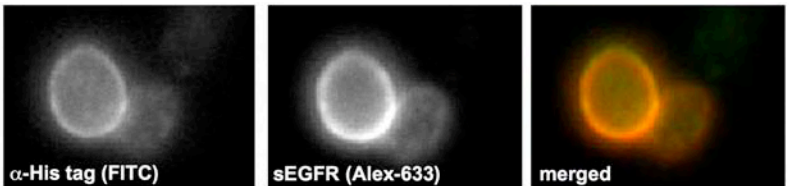


**Figure 11**
Yeast expressing hEGF-conjugated Aga2p were stained with a FITC-conjugated antibody against the his-tag (left) and labeled sEGFR.

Analysis using flow cytometry also showed that yeast expressing hEGF or Spitz can be detected by this method with an anti-His tag antibody and the isolated extracellular domain of the relevant receptor (Alexa-633 labeled). In Fig. 12, cells expressing his-tagged hEGF fused to Aga2p gave a range of fluorescence intensities (reflecting heterogeneity in expression level) when treated with a FITC-conjugated antibody against the histidine tag (see Figure 12A). When an Alexa633-labeled form of the extracellular domain from human EGFR was also added, it can be seen in Fig. 12B that cells with the highest FITC signal (*i.e.* with the largest numbers of cell-surface his tags) also bind most strongly to sEGFR. These data show both that the surface-displayed hEGF is active in binding to its receptor and that FACS can be used to monitor this system. Fig. 12C shows similar results for the Drosophila EGFR ligand Spitz. Cells expressing the largest amount of his-tagged Spitz were also sortable in FACS based on their binding to the Alexa633-labeled extracellular domain of the Drosophila EGF receptor.

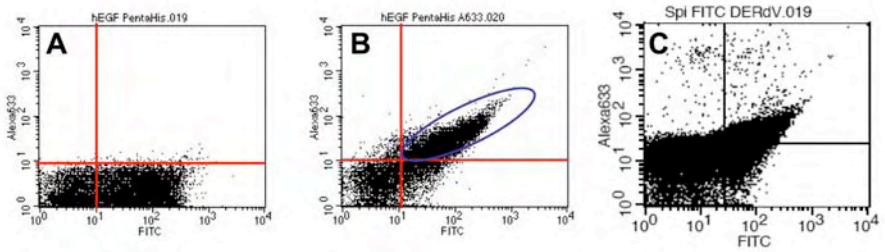


**Figure 12**
FACS analysis of yeast expressing Aga2p fused to hEGF (A and B) or Spitz (C). See text for details.

Having established that we can express hEGF and Spitz on the surface of EBY100 yeast, we proceeded to do the same for Argos, and for $Argos_{217}$. Cell surface-displayed $Argos_{217}$ was readily detectable using flow cytometry or immunofluorescence microscopy with an antibody against its histidine tag (Fig. 13). We could also detected cell-surface $Argos_{217}$ using Alexa633-labeled Spitz. Thus, we took this project to the stage at which we were in a position to screen libraries of Argos variants.

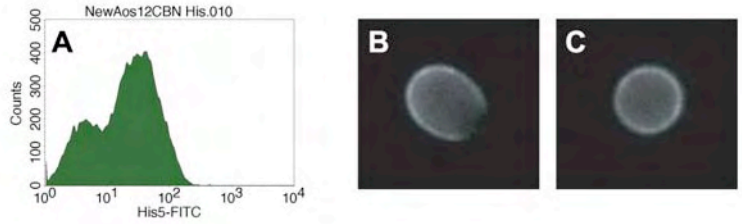


**Figure 13**
Surface expression of Aga2p-fused $Argos_{217}$ in yeast.
(**A**) the majority of expressing cells are detectable with anti-His antibody. (B and C) show immunofluorescence staining of surface Argos217 with α-His antibody.

We reached this point at around the same time that our structural studies pointed out that hDkks are likely to have no structural resemblance to Argos. Isolating a variant of Drosophila Argos that can bind human EGF-like ligands would be academically interesting and useful, but would have limited (or non-existent) clinical utility. We therefore chose to pause this project – in part because structure determination took a great deal of manpower – until the structures were

completed. It then became apparent that there are three human proteins: uPAR, the C4.4A homologue, and CD177/PRV-1 that should be subjected to the analysis described here. These studies are now in progress, and variants of these human proteins that can bind strongly to human ErbB ligands could be protein therapeutics of great value that represent precisely what our proposed research aimed to achieve.

Task 2c *Define conditions under which the observed weak binding of Argos-bearing phage to human EGF can be readily detected - to ensure that some signal can be obtained in f (below), and that this can be 'tuned' by altering stringency.*

This task was not completed, but will be undertaken in studies of uPAR, the C4.4A homologue, and CD177/PRV-1.

Task 2d *Generate mutants for display on M13 to determine which regions of hDkk1/2 and Argos are primarily responsible for defining the specificity of EGF domain recognition, and test these mutants by phage ELISA (months 8 to 12)*

Rather than using phage or yeast display to address this question, we took advantage of a recent collaboration in our studies of Argos to identify critical parts of the EGF domain binding site on Argos *in vivo*. We collaborated with Joseph Duffy at Indiana University, who undertook an ethyl-methane sulfonate (EMS) mutagenesis screen designed to recover modifiers of an Argos misexpression phenotype in the developing Drosophila eye [28-30]. This screen generated suppressors (loss-of-function alleles) and enhancers (gain-of-function) of the Argos misexpression phenotype, which can correspond to mutations in the *argos* transgene (intragenic) or in other genes (extragenic). Following mapping of these modifiers (suppressors or enhancers) to a single gene, the responsible lesions were identified by sequencing, utilizing non-mutagenized parental flies as a wild-type reference. This screen led to the identification of an allelic series of mutations in the *argos* transgene that decrease its activity.

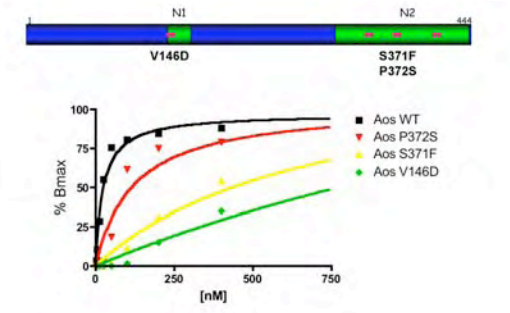


**Figure 14**
In the upper panel, a schematic representation of Argos sequence is shown. Pink circles represent sites at which mutations cause loss-of-function in the Duffy lab misexpression screen. In the lower panel, Biacore-derived curves are shown for binding of wild-type, P372S, S371F, and V146D Argos to immobilized Spitz. The S371F and V146D mutations significantly impair Spitz binding, and the P372S mutation has a moderate effect. Thus, loss-of-function in this *in vivo* screen correlates with impaired Spitz binding *in vitro*.

Critically, in addition to mutations at conserved cysteines (which probably cause misfolding – and thus loss of function – of Argos), two mutations in the C-terminal region of Argos (S371F and P372S), and one within the N-terminal cysteine-rich cluster domain (V146D) were identified. We produced recombinant protein corresponding to the V146D, S371F, and P372S mutants of Argos by secretion from baculovirus-infected Sf9 cells, and used Biacore to assess the binding of these mutated proteins to immobilized Spitz (Fig. 14). None of the mutations adversely affected protein production, arguing that they do not substantially affect folding or stability.

Mutation of V146 or S371 had a very significant effect on Spitz binding, reducing affinity by 46-fold and 37-fold respectively (Fig. 14). The P372S mutation, which also resulted in a loss of function in the Argos misexpression screen, reduced Spitz binding by around 2.5-fold, from a $K_D$ of 15nM for wild-type Argos to 32nM for P372S Argos. These results actually represent the first correlation between Argos function *in vivo* and its ability to bind and sequester Spitz. They also

identified the loop containing S371 and P372 (see Fig. 14) and that containing V146 as regions that would have been useful to target in our first round of sequence randomization for phage display in the coming year. Our structural studies [2] provided a clear rationale for the effects of these mutations. These data were published in a collaborative *J. Biol. Chem.* paper in 2006 [17] that is included in the Appendix.

Task 2e *Design and generate phage libraries in which sets of 6 amino acids in a region defined in d above (and structural studies of Argos/Spitz complex) are randomized (months 12 to 16)*

This task was not completed, but will be undertaken in studies of uPAR, the C4.4A homologue, and CD177/PRV-1.

Task 2f *Screen phage libraries at a range of different 'stringencies' (defined in c above) to isolate members with improved affinity for immobilized human EGF (months 16 to 24)*

This task was not completed, but will be undertaken in studies of uPAR, the C4.4A homologue, and CD177/PRV-1.

Task 2g *Analyze results of phage display analysis, and design molecules that combine the selected elements in all of the loops randomized. Test these for EGF binding (months 24 to 26).*

This task was not completed, but will be undertaken in studies of uPAR, the C4.4A homologue, and CD177/PRV-1.

Task 2h *Generate recombinant proteins incorporating the sequences indicated by phage display, and test for binding to human EGF and inhibition of EGF-induced ErbB receptor activation (months 26 to 30)*

This task was not completed, but will be undertaken in studies of uPAR, the C4.4A homologue, and CD177/PRV-1.

Task 2i *Pursue similar strategies for generating molecules capable of binding to other ErbB ligands (months 30-36 and beyond)*

This task was not completed, but will be undertaken in studies of uPAR, the C4.4A homologue, and CD177/PRV-1.

Task 2j *Develop strategies with collaborators for further testing of human Argos equivalents in more physiological settings.*

We still hope to initiate these efforts; it was always clear that they would not be achieved within the 3 year period of this IDEA Award. Our proposed studies with the C4.4A homologue, and CD177/PRV-1 are highly likely to lead us to the human Argos analog, which we believe will represent an excellent lead for future protein therapeutics that target the EGFR axis in breast cancer.

## KEY RESEARCH ACCOMPLISHMENTS

- Produced recombinant protein for hDkk1 and hDkk3
- Determined that hDkk1 and hDkk3 do not bind human ErbB ligands – arguing that the first hypothesis to be tested in our proposed research is incorrect. This defines the direction of our studies to generate novel anti-EGFs based on Dkk and Argos structural scaffolds
- Identified smallest fragment of Argos ($Argos_{217}$) that maintains function
- Crystallized $Argos_{217}$ for X-ray crystallographic studies
- Identified two inter-cysteine loops in Argos that are critical for EGF domain binding, allowing us to focus on these regions for randomization in phage display studies (published a paper on this)
- Crystallized an $Argos_{217}$/Spitz complex, and collected complete native data
- Using halide-soaked crystals, solved the structure of the $Argos_{217}$/Spitz complex to 1.6Å resolution (PDB code 3C9A)
- Solved the $Argos_{217}$ crystal structure to 2.5Å resolution by molecular replacement (PDB code 3CGU)
- Solved the structure of Spitz to 1.5Å resolution by molecular replacement (PDB code 3CA7)
- Determined that we cannot display function Argos on the surface of phage M13, requiring a change in experimental strategy
- Replacing phage display, we established the yeast surface display system in the laboratory, and showed that we can display active EGF, Spitz, and Argos on the surface of *S. cerevisiae*
- Structural studies identified the ligand-binding domain of the TGFβ receptor as an unexpected structural homolog of each Argos domain.
- Identified the unexpected structural similarity between Argos and uPAR, leading in turn to the hypothesis that the C4.4A homologue, and CD177/PRV-1 could possibly represent human analogs of Argos. This will be tested in the near future.

## REPORTABLE OUTCOMES

1. In 2005, results were presented in invited talks at ASBMB Meeting in San Diego, CA (April 2005) and FASEB Meeting on Receptors and Signal Transduction, Snowmass, CO (August 2005).

2. In 2006, results were presented in invited talks at a meeting of the Ludwig Institute for Cancer Research (May, 2006), a meeting on Protein Phosphorylation and Cell Signaling at the Salk Institute (August 2006), and NIDDK (December 2006).

3. In 2007, results were presented in invited talks at Vanderbilt University (January 2007), at the ASBMB Meeting in Washington DC (April 2007), at University of Colorado Health Sciences Center (May 2007), and at a meeting of the 'Receptor Tyrosine Kinase Consortium' in Cambridge, UK (July 2007).

4. In 2008, results were presented in invited talks at Einstein Medical College (January 2008), at New York University Medical Center (February 2008), at University College, London (March 2008), at the UK's National Institute for Medical Research (March 2008), and at NIEHS in Raleigh/Durham (April 2008).

5. In 2006, a paper was published describing identification of mutated forms of Argos that no longer bind Spitz:
D. Alvarado, T.A. Evans, R. Sharma, M.A. Lemmon, and J.B. Duffy. (2006). "*Argos mutants define an affinity-based threshold for Spitz inhibition in vivo*", *J. Biol. Chem.* **281,** 28993-29001. (see Appendix)

6. In 2008, we published a paper in *Nature* describing the Argos structure and its complex with Spitz (see Appendix):
D.E. Klein, S.E. Stayrook, F. Shi, K. Narayan and M.A. Lemmon. (2008). "*Structural basis for EGFR ligand sequestration by Argos*", *Nature*, published online, May 25, 2008 (doi:10.1038/nature06978).

7. Filed provisional patent application: Lemmon, M.A., Klein, D.E., and Stayrook, S.E.: EGF-Argos complex crystal and use thereof to identify inhibitors of EGFR signaling. U.S. Serial No: 61/064,135, Filed 19 Feb, 2008

8. Daryl Klein was awarded his Ph.D. in January of 2008 based on his contribution to this work. Dr. Klein also won a prestigious 2008 Harold M. Weintraub Graduate Student Award for his Ph.D. work (http://www.fhcrc.org/science/basic/weintraub/recipients.html#2008).

9. Secured NIH R01 funding for future aspects of this project: NIH grant 1R01CA125432-01, "Mechanisms of invertebrate EGF receptor inhibition", which will be funded until 05/31/2012

## CONCLUSIONS

The key finding during this project is embodied in our crystal structures of Argos alone, and Argos bound to Spitz. These structures reveal the structural basis for sequestration of EGF receptor ligands by Argos. Argos does not contain an EGF domain as previously suggested and supposed, and is also structurally unrelated to the EGF receptor. Instead, Argos unexpectedly contains three closely-related disulphide-bonded domains that are structurally similar to the ligand binding regions of TGFβ and uPA receptors, and were not recognized by sequence analysis. This structural similarity invalidates our previous supposition that the human Dickkopf proteins may be related to Argos. However, we are very excited by the fact that the structure identifies several candidate human Argos analogues in uPAR, the C4.4A analogue, and CD177/PRV-1. The three domains of Argos form a structure reminiscent of a C-clamp that neutralizes EGF-related growth factors through a bivalent capture mechanism that resembles the mode of ligand binding used by the EGF receptor itself. The uPA receptor does the same thing for an EGF domain in its ligand (uPA). We hypothesize (and will test) that C4.4A or CD177/PRV-1 similarly bind EGF domains, and the EGF domains that they recognize might be ErbB signaling molecules.

Simultaneously, we have developed a convenient yeast-based library display approach that will be useful for selecting C4.4A, uPAR or CD177/PRV-1 mutants if none binds human ErbB ligands. This method will be crucial if we find that the next step is to use one of these unexpected structural analogs of Argos as a 'scaffold' for a novel ligand-binding protein.

By either discovering that C4.4A or CD177/PRV-1 bind ErbB ligands, or by screening for randomly mutated versions of these proteins that will, we will have succeeded in our primary goal of identifying a human protein with Argos-like properties than can ultimately be developed into a protein therapeutic.

### So What ?

It is increasingly clear that excessive or unregulated expression (or shedding) of ErbB family ligands is important in numerous cancers, through autocrine and/or paracrine activation of cell growth. The role of ErbB ligands may be particularly important in cancers where available receptor-targeted approaches have failed or have met resistance – as in breast cancer, where EGFR-targeted agents do not appear to be very useful. In these (and other) cases, therapeutic agents that neutralize ErbB receptor ligands are likely to have great value. The understanding of Spitz neutralization by Argos that we have uncovered provides new avenues to explore in efforts to identify a human homologue of Argos. On one hand, we have frankly identified candidate proteins that might fulfil this function in humans (C4.4A and CD177/PRV-1). If they do, then these could be developed as protein therapeutics rather rapidly. If these proteins do not bind human ErbB ligands, then our structural lessons provide clear suggestions for how they might be used as scaffolds for generating new protein therapeutics that sequester aberrantly-produced EGF-like growth factors. We have established methods for doing this. Through either the direct or direct route, we anticipate that our work will ultimately lead to the development of ligand-targeted therapeutics based on human Argos-like molecules – exploiting a mechanism for inhibiting EGFR signalling that has evolved naturally.

## References


1 Klein, D. E., Nappi, V. M., Reeves, G. T., Shvartsman, S. Y. and Lemmon, M. A. (2004) Argos inhibits epidermal growth factor receptor signalling by ligand sequestration. Nature. **430**, 1040-1044

2 Klein, D. E., Stayrook, S. E., Shi, F., Narayan, K. and Lemmon, M. A. (2008) Structural basis for EGFR ligand sequestration by Argos. Nature, May 25. [Epub ahead of print]

3 Slamon, D. J., Leyland-Jones, B., Shak, S., Fuchs, H., Paton, V., Bajamonde, A., Fleming, T., Eiermann, W., Wolter, J., Pegram, M., Baselga, J. and Norton, L. (2001) Use of chemotherapy plus a monoclonal antibody against HER2 for metastatic breast cancer that overexpresses HER2. New Engl. J. Med. **344**, 783-792

4 Tuma, R. S. (2007) Lapatinib moves forward in inflammatory and early HER2-positive breast cancer trials. J. Natl. Cancer Inst. **99**, 348-349

5 Slamon, D. J., Goldolphin, W., Jones, L. A., Holt, J. A., Wong, S. G., Keith, D. E., Levin, W. J., Stuart, S. G., Udove, J., Ullrich, A. and Press, M. F. (1989) Studies of the HER-2/neu proto-oncogene in human breast and ovarian cancer. Science. **244**, 707-712

6 Klijn, J. G., Berns, P. M., Schmitz, P. I. and Foekens, J. A. (1992) The clinical significance of epidermal growth factor receptor (EGF-R) in human breast cancer: a review on 5232 patients. Endocr. Rev. **13**, 3-17

7 Mendelsohn, J. and Baselga, J. (2000) The EGF receptor family as targets for cancer therapy. Oncogene. **19**, 6550-6565

8 Fry, D. W. (2003) Mechanism of action of erbB tyrosine kinase inhibitors. Exp. Cell Res. **284**, 131-139

9 Ranson, M. and Sliwkowski, M. X. (2002) Perspectives on anti-HER monoclonal antibodies. Oncology. **63**, 17-24

10 Arteaga, C. L. and Truica, C. I. (2004) Challenges in the development of anti-epidermal growth factor receptor therapies in breast cancer. Semin. Oncol. **31(1 Suppl 3)**, 3-8

11 Dancey, J. E. and Freidlin, B. (2003) Targeting epidermal growth factor receptor - are we missing the mark ? Lancet. **362**, 62-64

12 Mendelsohn, J. and Baselga, J. (2003) Status of epidermal growth factor receptor antagonists in the biology and treatment of cancer. J. Clin. Oncol. **21**, 2787-2799

13 Di Fiore, P. P., Pierce, J. H., Kraus, M. H., Segatto, O., King, C. R. and Aaronson, S. A. (1987) ErbB-2 is a potent oncogene when overexpressed in NIH/3T3 cells. Science. **237**, 178-182

14 Di Fiore, P. P., Pierce, J. H., Fleming, T. P., Hazan, R., Ullrich, A., King, C. R., Schlessinger, J. and Aaronson, S. A. (1987) Overexpression of the human EGF receptor confers an EGF-dependent transformed phenotype to NIH 3T3 cells. Cell. **51**, 1063-1070

15 Rösel, M., Claas, C., Seiter, S., Herlevsen, M. and Zöller, M. (1998) Cloning and functional characterization of a new phosphatidyl-inositol anchored molecule of a metastasizing rat pancreatic tumor. Oncogene. **17**, 1989-2002

16 Temerinac, S., Klippel, S., Strunck, E., Röder, S., Lübbert, M., Lange, W., Azemar, M., Meinhardt, G., Schaefer, H. E. and Pahl, H. L. (2000) Cloning of PRV-1, a novel member of the uPAR receptor superfamily, which is overexpressed in polycythemia rubra vera. Blood. **95**, 2569-2576

17 Alvarado, D., Evans, T. A., Sharma, R., Lemmon, M. A. and Duffy, J. B. (2006) Argos mutants define an affinity threshold for spitz inhibition in vivo. J. Biol. Chem. **281**, 28993-29001

18 Allendorph, G. P., Vale, W. W. and Choe, S. (2006) Structure of the ternary signaling complex of a TGF-beta superfamily member. Proc. Natl. Acad. Sci. U. S. A. **103**, 7643-7648

19 Greenwald, J., Fischer, W. H., Vale, W. W. and Choe, S. (1999) Three-finger toxin fold for the extracellular ligand-binding domain of the type II activin receptor serine kinase. Nat. Struct. Biol. **6**, 18-22

20 Barinka, C., Parry, G., Callahan, J., Shaw, D. E., Kuo, A., Bdeir, K., Cines, D. B., Mazar, A. and Lubkowski, J. (2006) Structural basis of interaction between urokinase-type plasminogen activator and its receptor. J. Mol. Biol. **363**, 482-495

21 Huai, Q., Mazar, A. P., Kuo, A., Parry, G. C., Shaw, D. E., Callahan, J., Li, Y., Yuan, C., Bian, C., Chen, L., Furie, B., Furie, B. C., Cines, D. B. and Huang, M. (2006) Structure of human urokinase plasminogen activator in complex with its receptor. Science. **311**, 656-659

22 Hansen, L. V., Laerum, O. D., Illemann, M., Nielsen, B. S. and Ploug, M. (2008) Altered expression of the urokinase receptor homologue, C4.4A, in invasive areas of human esophageal squamous cell carcinoma. Int. J. Cancer. **122**, 734-741

23 Allendorph, G. P., Vale, W. W. and Choe, S. (2006) Structure of the ternary signaling complex of a TGF-beta superfamily member. Proc. Natl. Acad. Sci. U. S. A. . **103**, 7643-7648

24 Chao, G., Lau, W. L., Hackel, B. J., Sazinsky, S. L., Lippow, S. M. and Wittrup, K. D. (2006) Isolating and engineering human antibodies using yeast surface display. Nat. Protoc. **1**, 755-768

25 Boder, E. T. and Wittrup, K. D. (1997) Yeast surface display for screening combinatorial polypeptide libraries. Nat. Biotechnol. **15**, 553-557

26 Kim, Y. S., Bhandari, R., Cochran, J. R., Kuriyan, J. and Wittrup, K. D. (2006) Directed evolution of the epidermal growth factor receptor extracellular domain for expression in yeast. Proteins. **62**, 1026-1035

27 Jin, M., Song, G., Carman, C. V., Kim, Y. S., Astrof, N. S., Shimaoka, M., Wittrup, D. K. and Springer, T. A. (2006) Directed evolution to probe protein allostery and integrin I domains of 200,000-fold higher affinity. Proc. Natl. Acad. Sci. U S A. **103**, 5758-5763

28 Alvarado, D., Rice, A. H. and Duffy, J. B. (2004) Bipartite inhibition of Drosophila epidermal growth factor receptor by the extracellular and transmembrane domains of Kekkon1. Genetics. **167**, 187-202

29 Guichard, A., Srinivasan, S., Zimm, G. and Bier, E. (2002) A screen for dominant mutations applied to components in the Drosophila EGF-R pathway. Proc. Natl. Acad. Sci. U S A. **99**, 3752-3757

30 Penton, A., Wodarz, A. and Nusse, R. (2002) A mutational analysis of dishevelled in Drosophila defines novel domains in the dishevelled protein as well as novel suppressing alleles of axin. Genetics. **161**, 747-762

**Bibliography of all publications and meeting abstracts**

*i.* Research publications:

1. D. Alvarado, T.A. Evans, R. Sharma, M.A. Lemmon, and J.B. Duffy. (2006). "*Argos mutants define an affinity-based threshold for Spitz inhibition in vivo*", *J. Biol. Chem.* **281,** 28993-29001.

2. D.E. Klein, S.E. Stayrook, F. Shi, K. Narayan and M.A. Lemmon. (2008). "*Structural basis for EGFR ligand sequestration by Argos*", *Nature*, published online, May 25, 2008 (doi:10.1038/nature06978).

*ii.* Published meeting abstracts:

1. Lemmon, M.A. (2007). Activation and inhibition of the EGF receptor. *FASEB Journal* **21,** Issue: 5 Pages: A46-A46 Published: Apr 2007

2. Lemmon, M.A. (2005). Activation and inhibition of the EGF receptor. *FASEB Journal* **19,** Issue: 5, Part 2 Suppl. Pages: A1690-A1690 Published: Mar 7, 2005

**List of personnel receiving pay from the research effort:**

Mark A. Lemmon, Ph.D.
Pamela Burgess-Jones
Daryl E. Klein

## Appendix

**Two publications are attached:**

**1.** D. Alvarado, T.A. Evans, R. Sharma, M.A. Lemmon, and J.B. Duffy. (2006). "*Argos mutants define an affinity-based threshold for Spitz inhibition in vivo*", *J. Biol. Chem.* **281,** 28993-29001. Attached as pdf reprint.

2. D.E. Klein, S.E. Stayrook, F. Shi, K. Narayan and M.A. Lemmon. (2008). "*Structural basis for EGFR ligand sequestration by Argos*", *Nature*, published online, May 25, 2008 (doi:10.1038/nature06978).
The final accepted version of this manuscript is attached (not yet in *Nature* format), plus Supplementary information. The final paper will be published in the June 26 2008 issue of *Nature*.

THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 281, NO. 39, pp. 28993–29001, September 29, 2006


# Argos Mutants Define an Affinity Threshold for Spitz Inhibition *in Vivo*[*]

Received for publication, April 20, 2006, and in revised form, July 24, 2006 Published, JBC Papers in Press, July 26, 2006, DOI 10.1074/jbc.M603782200

**Diego Alvarado**[‡1], **Timothy A. Evans**[§2], **Raghav Sharma**[§3], **Mark A. Lemmon**[‡], **and Joseph B. Duffy**[§4]

*From the* [‡]*Department of Biochemistry and Biophysics, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania 19104-6059 and the* [§]*Department of Biology, Indiana University, Bloomington, Indiana 47405*

**Argos, a secreted antagonist of *Drosophila* epidermal growth factor receptor (dEGFR) signaling, acts by sequestering the activating ligand Spitz. To understand how different domains in Argos contribute to efficient Spitz sequestration, we performed a genetic screen aimed at uncovering modifiers of an Argos misexpression phenotype in the developing eye. We identified a series of suppressors mapping to the Argos transgene that affect its activity in multiple developmental contexts. These point mutations map to both the N- and C-terminal cysteine-rich regions, implicating both domains in Argos function. We show by surface plasmon resonance that these Argos mutants are deficient in their ability to bind Spitz *in vitro*. Our data indicate that a mere ~2-fold decrease in $K_D$ is sufficient to compromise Argos activity *in vivo*. This effect could be recapitulated in a cell-based assay, where a higher molar concentration of mutant Argos was needed to inhibit Spitz-dependent dEGFR phosphorylation. In contrast, a ~37-fold decrease in the binding constant nearly abolishes Argos activity *in vivo* and in cellular assays. In agreement with previously reported computational studies, our results define an affinity threshold for optimal Argos inhibition of dEGFR signaling during development.**

The epidermal growth factor receptor belongs to a family of receptor tyrosine kinases that are well conserved from lower metazoans to humans (1, 2). In humans, mutations or genetic alterations that alter receptor activity have been correlated with cancer progression and poor clinical outcome, validating the ErbB family as a target for therapeutic agents (3–7).

In *Drosophila melanogaster*, EGF[5] receptor signaling is utilized reiteratively throughout development to mediate a wide array of cellular decisions (8). Remarkably, this versatility is accomplished with a single receptor (dEGFR) and four activating ligands: Spitz (Spi), Gurken, Keren, and Vein. The functional diversity of dEGFR signaling has been partially attributed to the differential use of ligands throughout development (8). For instance, the ligand Gurken is produced exclusively in the germline to specify eggshell structures and the embryonic axes (9, 10). In contrast, Spi and Vein participate in numerous processes, sometimes sharing additive roles (such as in ventral ectoderm patterning) (11–14) and in some instances acting as the main dEGFR ligand (such as in photoreceptor recruitment or wing vein differentiation, respectively) (15–20). Interestingly, no knock-out phenotype or expression pattern has been reported for the fourth ligand, Keren (21). In addition to the four agonists, the *Drosophila* EGF receptor signaling system includes two extracellular inhibitors, which function in negative feedback loops to antagonize dEGFR signaling. Kekkon 1 is a transmembrane molecule of the leucine-rich repeat-immunoglobulin (LIG) superfamily that attenuates dEGFR activity via a direct interaction (22, 23). Argos (Aos) is a secreted molecule that was initially proposed to bind and inhibit dEGFR by virtue of its atypical EGF domain (24). Recent work, however, has demonstrated that Aos instead exerts its antagonistic effect on dEGFR signaling by sequestering the activating ligand Spi, although its effect on the remaining three ligands has not yet been reported (25). Aos was also shown to associate with the surface of cultured S2 cells in an interaction that is dissociable with excess soluble heparin (25), suggesting a potential regulatory mechanism for Aos activity *in vivo*. However, the functional and physiological significance of this finding remains unclear.

Structurally, Aos is composed of N- and C-terminal cysteine-rich regions (NCR and CCR, respectively), separated by a largely unconserved linker. The NCR, which contains 4 cysteines, has no known direct function, and misexpression of this domain alone displays no biological activity. The CCR includes 12 cysteines (see Fig. 1*B*) and contains a putative EGF-like domain (residues 363–424) (24). Misexpression of the entire CCR from Aos (residues 225–444) displays partial activity *in vivo* and is sufficient for binding Spi *in vitro* (albeit with a ~20-fold decreased affinity) (25, 26). In contrast, misexpression of the putative EGF-like domain alone does not rescue Aos mutant phenotypes (26), arguing that (if it does adopt an EGF-like fold) it is not sufficient for Aos function.

Aos participates in multiple developmental processes where it is expressed as a "high threshold" gene, in response to high levels of Spi-induced dEGFR signaling (27, 28). Consistent with its role as an inhibitor of dEGFR signaling, Aos knock-outs exhibit phenotypes typical of dEGFR gain-of-function mutants,

* This work was supported by National Science Foundation Research Grant NSF IBN-0131707 (to J. B. D.), National Institutes of Health Grant RO1-CA079992 (to M. A. L.), and Department of Defense Breast Cancer Research Program Grant W81XWH-05-1-0289 (to M. A. L.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[1] Damon Runyon Fellow supported by Damon Runyon Cancer Research Fellowship DRG-1884-05. To whom correspondence may be addressed. Tel.: 215-898-3411; E-mail: dalvarad@mail.med.upenn.edu.

[2] Supported by National Institutes of Health Predoctoral Training Grant GM-007757.

[3] Recipient of an Howard Hughes Medical Institute Capstone Award.

[4] To whom correspondence may be addressed. Present address: Dept. of Biology and Biotechnology, Worcester Polytechnic Institute, Worcester, MA 01609. Tel.: 508-831-5579; E-mail: jduffy@wpi.edu.

[5] The abbreviations used are: EGF, epidermal growth factor; dEGFR, *Drosophila* EGF receptor; Aos, Argos; NCR, N-terminal cysteine-rich region; CCR, C-terminal cysteine-rich region; Spi, Spitz; GMR, glass multiple reporter; SPR, surface plasmon resonance; UAS, upstream activating sequence.

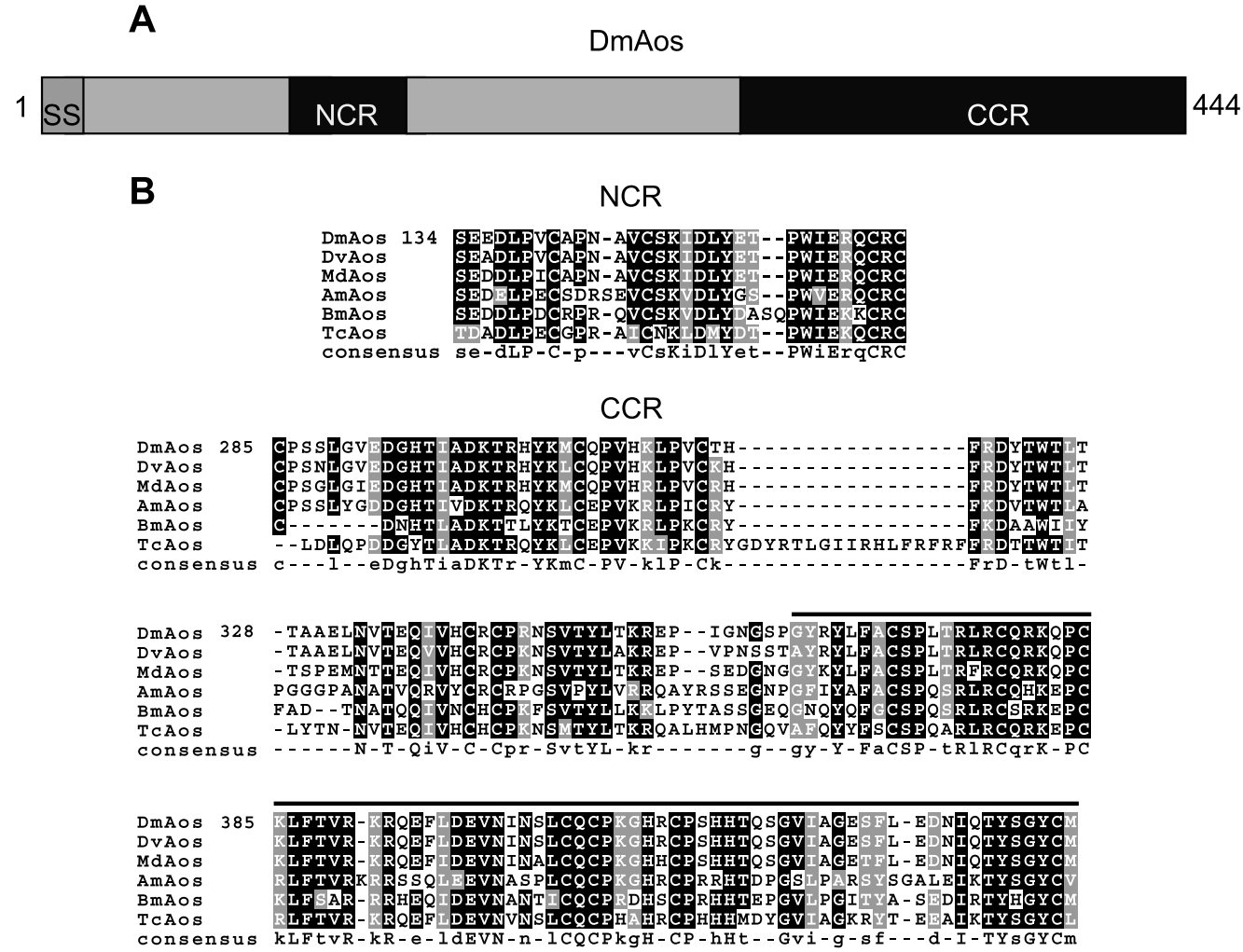


FIGURE 1. **Identification of Aos orthologs reveals two conserved regions.** *A*, graphic representation of *D. melanogaster* Aos depicting the relative placement and sizes of the NCR and the CCR. *B*, alignment of the NCR and CCR in Aos orthologs from two *Drosophila* species (*D. melanogaster*, *Dm*; *Drosophila virilis*, *Dv*), the house fly (*M. domestica*, *Md*) the honeybee (*A. mellifera*, *Am*), the silk moth (*Bombyx mori*, *Bm*), and the beetle (*T. castaneum*, *Tc*). Consensus residues are shown below with identities shaded in *black* and similarities marked in *gray*, and the putative EGF motif marked with a *black line*.

and Aos misexpression inhibits dEGFR signaling (26, 29). Aos contributes importantly to the spatio-temporal regulation of dEGFR signaling through its participation in a negative feedback loop as a result of Spi-dependent dEGFR signaling. For example, Aos is required for the proper timing of pulsations during oenocyte deamination (30). Aos has also been described as a long range inhibitor during eye development (and other tissues), acting to ensure the formation of steep Spi gradients close to the source of Spi production (31). Recent computational studies have proposed two key roles for Aos in dEGFR signaling based on a model of the dEGFR/Spi/Aos module in embryonic ventral ectoderm patterning (32). First, sequestration by Aos limits the spatial range of Spi action. Second, the Aos negative feedback loop counteracts fluctuations in gene dosage (Spi secretion rate and dEGFR levels), imparting robustness to the system (32). One important prediction of the model was that Spi sequestration by Aos must be nearly irreversible (or of very high affinity) to provide a robust feedback loop. As such, an increase in the off rate would result in a loss of robustness, although this remains to be tested experimentally.

To test the prediction that even small reductions in Spi/Aos affinity cannot be tolerated and to investigate the domain requirements for Aos activity, we screened for mutations in an Aos transgene that suppress the strong misexpression phenotype caused in the developing eye (29). We report the identification of a series of point mutations in Aos that reduce or impair its activity in multiple tissues. These lesions map to both the NCR and CCR regions, demonstrating that both modules are necessary for Aos function *in vivo*. We have also correlated reductions in the phenotypic strength of Aos mutants with decreases in their *in vitro* binding affinity for Spi. Whereas a mere ~2-fold reduction in affinity appears to be sufficient to reduce the effectiveness of Aos as a dEGFR inhibitor, a ~37-fold decrease in affinity greatly compromises its activity. We also show that these mutants are correspondingly less efficient in abolishing Spi-dependent dEGFR phosphorylation in cellular studies. Our data thus show that both the NCR and CCR in Aos are necessary for establishing a high affinity complex with Spi, which is critical for imparting dEGFR signaling robustness.

## EXPERIMENTAL PROCEDURES

*Identification of Aos Orthologs*—Publicly available genome sequences were searched using the tblastn algorithm for se-

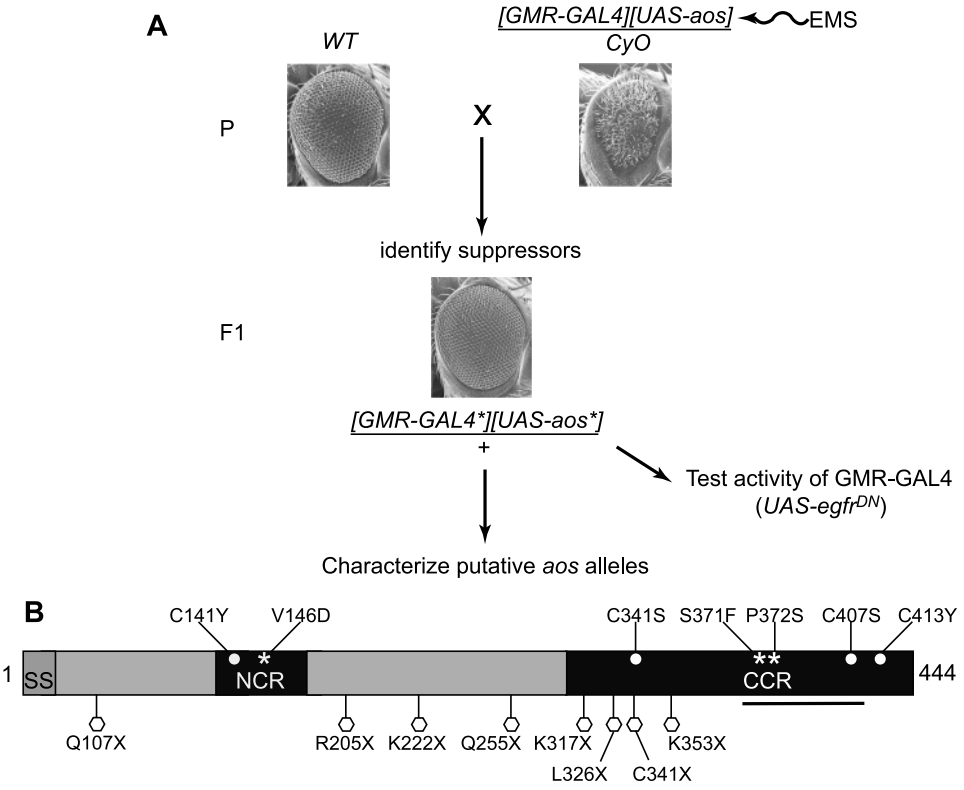


FIGURE 2. **A screen to identify mutations affecting Aos activity.** *A*, a strain exhibiting a rough eye phenotype caused by misexpression of Aos in the developing eye (GMR>Aos) was constructed and maintained over a balancer chromosome (*CyO*). This strain was ethyl methane sulfonate-mutagenized and outcrossed to a wild type strain. Progeny misexpressing Aos were then selected by the absence of the dominant curly wing phenotype associated with CyO balancer and then screened for suppression of the eye phenotype. Refer to the text and "Experimental Procedures " for additional details. *B*, graphic view of Aos from *D. melanogaster* depicting the signal sequence (*SS*), the two conserved regions, NCR and CCR, in *black*, the variable regions in *gray*, and the putative EGF domain marked with a *black line*. Class I mutants (non-cysteine missense) are shown as *asterisks*; class II mutants (cysteine missense) are shown as *dots*; and class III alleles (nonsense) are depicted with *hexagons*.

quences related to Aos. Sequence alignments were performed with ClustalW1.8 (searchlauncher.bcm.tmc.edu/multi-align/multi-align.html) and prepared for publication using the BOXSHADE server (www.ch.embnet.org/software/BOX_form.html).

*Genetics—P{GAL4-ninaE.GMR}, P{UAS-aos}/CyO* flies were generated by standard recombination methods from the individual P insertions. The males were mutagenized with 25 mM EMS (33) and crossed to *w; iso2; iso3* females at 27 °C. *P{GAL4-ninaE.GMR}, P{UAS-aos}* progeny were screened for suppression of the rough eye phenotype (see Fig. 2). Suppressors affecting activity of *P{GAL4-ninaE.GMR}* were identified by crossing them to *P{UAS-egfr$^{DN}$}/CyO* and were subsequently discarded. Suppressors that retained GAL4 activity were balanced and characterized further by mapping, sequencing, Western analysis, and misexpression with embryonic and wing drivers (*Tubulin-GAL4* and *MS1096-GAL4*). For embryonic activity, suppressor lines (*UAS-aos*/CyO*) were crossed to the *Tubulin-GAL4/CyO* strain, and percentage of viability was calculated as 100× (# flies with *straight wing*)/.5(*# flies with curly wing*).

For rescue analysis, a *sev-GAL4* driver was combined with each of the represented *UAS-aos* loss of function alleles (see Fig. 5) and then introduced into an *aos* null background (*aos$^{\Delta 7}$/Df(3L)Exel6129*). For scanning electron microscopy of the adult eye, the females were dehydrated in an increasing ethanol: $dH_2O$ series, as described by Tio *et al.* (18).

*Sequence and Western Blot Analysis of aos Alleles*—For each putative loss of function allele genomic DNA was isolated from 10–20 adult flies with Qiagen DNeasy columns (Qiagen). The *aos* transgene was then amplified by PCR, purified by gel extraction (Qiagen), and sequenced using cycle sequencing according to the manufacturer's instructions (Applied Biosystems). At least two independent rounds of genomic DNA purification, PCR, and sequencing were carried out for each allele. For Western analysis, ovaries from four females misexpressing the suppressors during stages 9–11 in the follicle cells (*CY2-GAL4, UAS-aos*) were dissected in phosphate-buffered saline, transferred into 50 μl of phosphate-buffered saline +25 μl of 4× sample buffer on ice, and homogenized. 15 μl of each sample was loaded on an 8% SDS-PAGE gel and transferred to nitrocellulose. The blots were probed with anti-Aos (Developmental Studies Hybridoma Bank) at 1:100, stripped, and reprobed with anti-α-tubulin (12G10; Developmental Studies Hybridoma Bank) at 1:5000 as a loading control.

*Molecular Cloning and Protein Production*—Full-length Aos was amplified by PCR, incorporating a SpeI restriction site in the 5′ primer and a $His_6$ tag followed by a NotI restriction site in the 3′ primer, and subcloned into pFastbac (Invitrogen) giving pFbAosHis. Aos mutants were generated by QuikChange (Stratagene) using pFb-AosHis as a template. Baculoviruses encoding different Aos alleles were generated and amplified according to the manufacturer's instructions. For protein purification, 1 liter of Sf9 cells were infected with each corresponding baculovirus (except Aos$^{V146D}$, which required 2–2.5 liters) for 3 days. Conditioned medium was dialyzed against 12 volumes of 10 mM HEPES, pH 8, 150 mM NaCl, and flowed through a nickel-nitrilotriacetic acid column with a 2-ml bed volume (Qiagen). The column was washed with 25 ml of 20 mM imidazole in buffer A (25 mM Tris, pH 8, 100 mM NaCl), and protein was eluted in 5 ml of 300 mM imidazole/buffer A. The eluted protein was concentrated and further purified by gel filtration using a Superose 6 column (Amersham Biosciences) equilibrated with 25 mM HEPES, pH 8, 150 mM NaCl. Secreted His-tagged Spi (amino acids 1–128) was purified from transfected S2 cells as described previously (25). The protein concentrations were determined using absorbance at 280 nm.

*Surface Plasmon Resonance (SPR)*—SPR experiments were performed on a BIAcore 3000. Spi was immobilized onto a CM5 sensorchip by standard amine coupling using 10 mM acetate, pH 5.5, for preloading the CM-dextran surface. Purified Aos

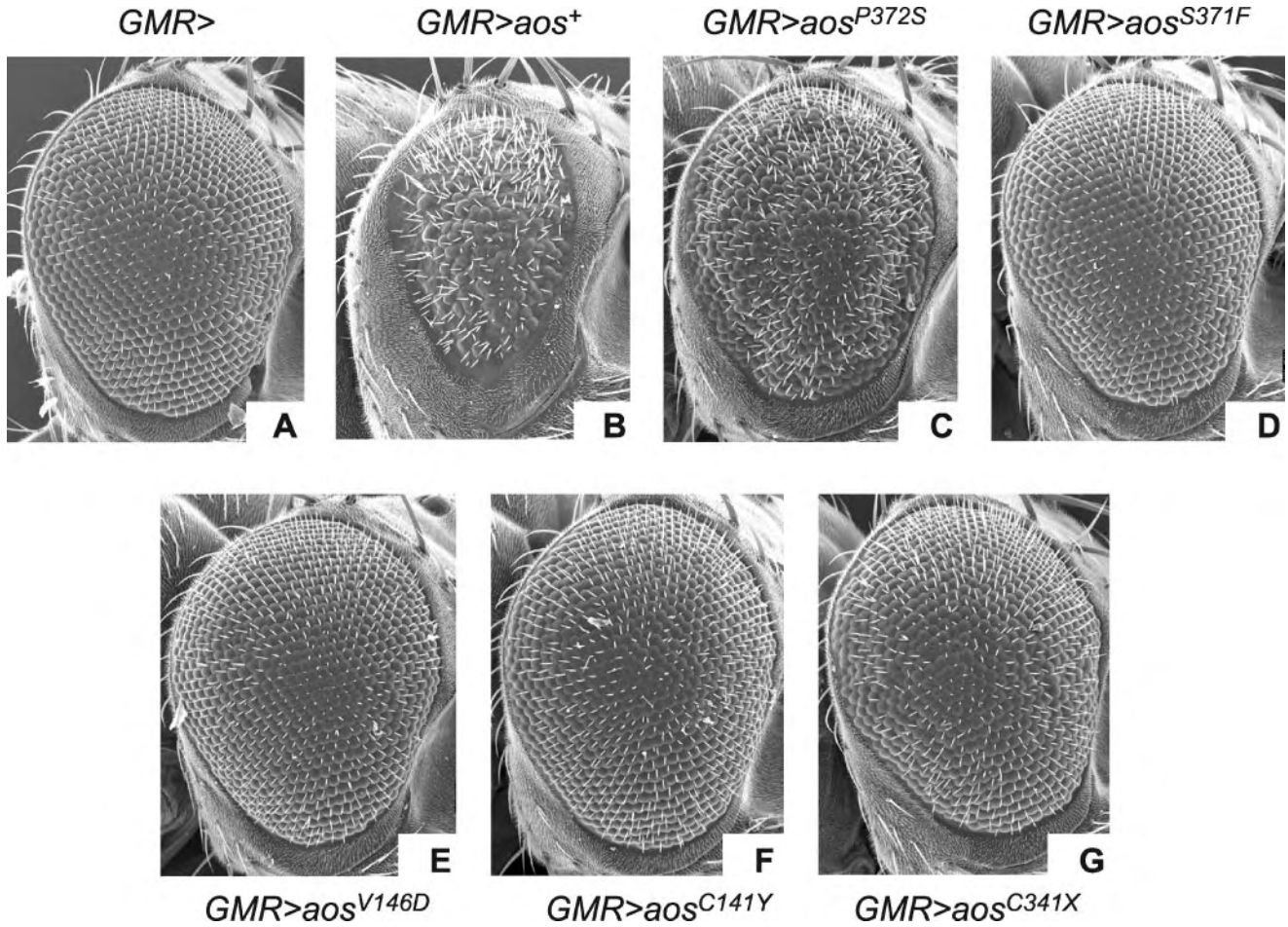


FIGURE 3. **Suppressors of Aos activity.** Scanning electron micrographs of adult compound eyes. *A* and *B* are from control flies expressing the GAL4 driver alone (*A*, *GMR-GAL4* only, GMR>) and the GAL4 driver with the nonmutagenized *aos* transgene (*B*, *GMR-GAL4 UAS-aos*$^+$, *GMR>aos*$^+$). *C–G* represent examples of moderate (*C*) and strong suppressors (*D–G*) of the *aos* misexpression phenotype.

mutants were flowed over the Spi-containing sensorchip from lower to higher concentrations at 10 $\mu$l/min. The sensorchip was regenerated after running each sample with 10 mM glycine, pH 3, and 1 M NaCl. Binding curves were generated by normalizing the total response units at equilibrium against the maximal saturation response ($B_{max}$) and plotting them as a function of protein concentration. Curves were fit to a single-site binding model using the program Prism, from which the $K_D$ values were derived. The experiments were done at least three times, generating error bars and standard error values. The experiments were also carried out by flowing protein from higher to lower concentration, yielding slightly higher $K_D$ values (likely because of sample aggregation and incomplete regeneration), but similar wild type to mutant ratios.

*Activation Studies*— dEGFR activation assays in S2 cells were performed as described previously (25). Briefly, dEGFR-expressing D2f cells (kindly provided by Benny Shilo) were serum-starved overnight, and dEGFR production was induced with 60 $\mu$M $CuSO_4$ for 3 h. The cells were incubated on ice with purified Spi alone or in the presence of the indicated amount of purified Aos proteins for 10 min. The cells were lysed in radioimmune precipitation assay buffer (25 mM Tris, pH 7.5, 150 mM NaCl, 1% Nonidet P-40, 1% sodium deoxycholate, and 0.1% SDS with phosphatase and protease inhibitors), and lysates were separated by SDS-PAGE and transferred to nitrocellulose. The blots were probed with anti-pY20 (Santa Cruz), stripped, and reprobed with anti-dEGFR as a loading control (23).

## RESULTS

*Isolation of Mutations Affecting Aos Activity*—Aos was first described in *D. melanogaster* as an antagonistic ligand of dEGFR with an atypical EGF motif that was hypothesized to direct its association with dEGFR and preclude Spi binding (24, 34). However, recent studies have shown that Aos inhibits dEGFR signaling by binding to the activating ligand Spi rather than associating with the receptor itself (25), bringing the significance of the atypical EGF-like domain in Aos into question. As pointed out by Howes *et al.* (26), sequences outside the putative EGF motif of Aos are strongly conserved in the house fly *Musca domestica* (representing an evolutionary distance of ~100 million years) and are required for Aos function *in vivo*, further suggesting that more than just the putative EGF-like domain is required for Aos function. We have expanded upon these observations by identifying Aos orthologs outside the dipteran lineage (Fig. 1). We found Aos orthologs in two additional arthropods, the honeybee *Apis mellifera* and the beetle *Tribolium castaneum*, indicating conservation of Aos over a span of ~500 million years. We were unable to identify Aos

**TABLE 1**
**Molecular and genetic characterization of *aos* mutants**

| Line | Nucleotide change | Amino acid change | Allelic class | Activity | | |
|---|---|---|---|---|---|---|
| | | | | Eye[a] | Wing[b] | Viability[c] |
| | | | | | | % |
| *UAS-aos*$^{+}$ | | | Wild type | ++ | 100 | 0 (124) |
| *UAS-aos*$^{39}$ | C1112T | S371F | Class I | − | 0 | 111 (522) |
| *UAS-aos*$^{136}$ | T437A | V146D | Class I | − | 0 | 107 (556) |
| *UAS-aos*$^{334}$ | C1114T | P372S | Class I | + | 4 | 10 (462) |
| *UAS-aos*$^{14}$ | T1021A | C341S | Class II | − | 0 | ND |
| *UAS-aos*$^{97}$ | G1238A | C413Y | Class II | − | 0 | ND |
| *UAS-aos*$^{134}$ | G422A | C141Y | Class II | − | 0 | ND |
| *UAS-aos*$^{154}$ | G422A | C141Y | Class II | − | 0 | 106 (410) |
| *UAS-aos*$^{182}$ | T1219A | C407S | Class II | − | 0 | ND |
| *UAS-aos*$^{9}$ | A949T | K317X | Class III | − | 0 | ND |
| *UAS-aos*$^{74}$ | T977A | L326X | Class III | − | 0 | ND |
| *UAS-aos*$^{142}$ | C319T | Q107X | Class III | − | 0 | ND |
| *UAS-aos*$^{160}$ | T1023A | C341X | Class III | − | 0 | 93 (511) |
| *UAS-aos*$^{222}$ | C763T | Q255X | Class III | − | 0 | ND |
| *UAS-aos*$^{233}$ | C613T | R205X | Class III | − | 0 | ND |
| *UAS-aos*$^{270}$ | A664T | K222X | Class III | − | 0 | ND |
| *UAS-aos*$^{325}$ | A1057T | K353X | Class III | − | 0 | ND |

[a] Activity based on the severity of the rough eye phenotype when the *UAS-aos* allele is combined with the *GMR-GAL4* driver. ++ indicates a severe rough eye phenotype, whereas + and − indicate moderate or weak phenotypes, respectively.
[b] Percentage of wings with gap in the L4 vein when the *UAS-aos* allele is combined with the *MS1096-GAL4* driver. The numbers refer to the percentages of times a gap in the L4 wing vein was observed.
[c] Expression of *aos* mutants was driven ubiquitously during embryogenesis using the *Tubulin-GAL4* driver. The percentage of viability was determined as described under "Experimental Procedures." The numbers in parentheses represent the total number of flies scored. ND, not determined.

orthologs in vertebrates, suggesting that, like the dEGFR inhibitor Kekkon1, the existence of Aos is phylogenetically restricted (35, 36). Sequence conservation in Aos extends well beyond the boundaries of the putative EGF motif and reveals two strongly conserved regions (Fig. 1) that we term the NCR and CCR, which are separated by a region of variable length. The NCR and CCR are defined primarily by sets of 4 and 12 conserved cysteines, respectively. Previous studies have indicated that the CCR represents the primary ligand-binding domain and can provide a measure of activity *in vivo* (25, 26). In contrast, no direct functionality has been ascribed to the NCR.

Given these regions of strong sequence conservation and the reported function of Aos as a "Spi sink," we sought to investigate the contributions of each domain to Aos function. We therefore carried out a screen for mutations that disrupt the ability of an Aos transgene to affect eye development upon misexpression, as diagrammed in Fig. 2*A*. During eye development Aos limits Spi availability, thus attenuating dEGFR signaling and preventing the specification of excess photoreceptors (24, 29, 31). Alterations in the dosage of *aos* (and consequently dEGFR activity) are readily observed as disruptions in the highly organized array of facets in the adult compound eye. Misexpression of an Aos transgene in the developing eye (using *GMR-GAL4*) leads to inhibition of dEGFR signaling (presumably through sequestration of Spi) and causes a severe loss of photoreceptors and morphological defects in the adult eye (Figs. 2*A* and 3*B*) (37–39). Reducing transgene activity would restore normal morphology to the eye and thus provides a means to identify mutations in Aos that impair its activity *in vivo*.

To carry out a screen for functionally defective *aos* transgenes, we utilized the GAL4/UAS system (37, 40). In this bipartite expression system, regulation of the transgene of interest, here *aos*, is controlled by UASs bound by the yeast transcriptional activator GAL4. *Drosophila* strains expressing GAL4 in specific spatio-temporal patterns are then used to direct UAS transgene expression in the tissue of interest. For this screen we generated a recombinant fly strain with the *UAS-aos* transgene under the control of a GAL4 line, P{*GAL4-ninaE.GMR*}, that drives GAL4 expression (and thus *Aos* misexpression) in the developing eye. This strain, referred to as *GMR>aos*$^{+}$, was mutagenized and outcrossed, and the progeny were screened for suppression of the eye defects associated with Aos misexpression (Fig. 2*A*). In addition to recovering mutations in the *aos* transgene, the primary focus of this work, we anticipated the recovery of two additional classes of mutations that would also suppress the *GMR>aos*$^{+}$ eye phenotype shown in Fig. 3*B*. Mutations in the GAL4 driver would prevent misexpression of the *aos* transgene and thus revert the associated eye phenotype. These mutations were identified as progeny that failed to show eye defects when combined with a distinct UAS responder transgene, here a dominant negative version of dEGFR (*UAS-dEGFR*$^{DN}$), and were discarded (Fig. 2*A*). Progeny that retained an intact *GMR-GAL4* driver (and thus Aos misexpression) were further characterized. Mutations in genes essential for Aos function might also be recovered and as such would be unlinked to the *UAS-aos* transgene (38, 39). Here we report the isolation and characterization of 16 mutations linked to the *aos* transgene that resulted in suppression of the misexpression phenotype in the eye (Table 1 and Figs. 2*B* and 3). Of these 16 lines, 15 showed strong suppression in the eye, consistent with minimal Aos function. The remaining mutant (*aos*$^{334}$, identified as Aos$^{P372S}$ below) showed only partial suppression in the eye (Table 1 and Fig. 3*C*), suggesting that it retained moderate activity. In addition to its function in the eye, Aos has been implicated in regulating dEGFR signaling in other contexts, including wing development and embryogenesis (13, 34). Suppressor activity was assessed in these two contexts as well, using wing-specific (*MS1096-GAL4*) and ubiquitous (*Tubulin-GAL4*) drivers. Misexpression of wild type *aos* with *MS1096-GAL4* leads to a loss of wing veins, whereas embryonic misexpression with *Tubulin-GAL4* results in lethality. All of the

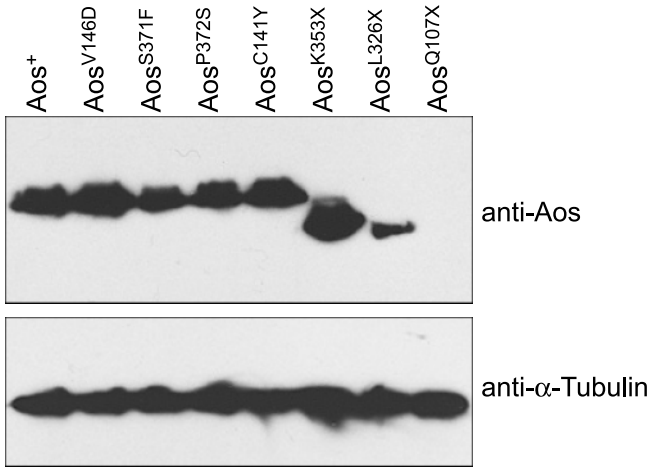


FIGURE 4. **Aos mutants are expressed at similar levels in fly ovaries.** Western blot analysis of the three class I mutations (*UAS-aos$^{V146D}$*, *UAS-aos$^{S371F}$*, and *UAS-aos$^{P372S}$*), a representative class II allele (*UAS-aos$^{C141Y}$*), and three class III mutations encoding respectively shorter isoforms (*UAS-aos$^{K353X}$*, *UAS-aos$^{L326X}$*, and *UAS-aos$^{Q107X}$*) is shown. Similar levels of expression for these alleles are observed upon misexpression, with the exception of *UAS-aos$^{Q107X}$*, which eliminates the epitope recognized by the Aos antibody (26). Anti-α-Tubulin loading controls are shown in the *bottom panel*.

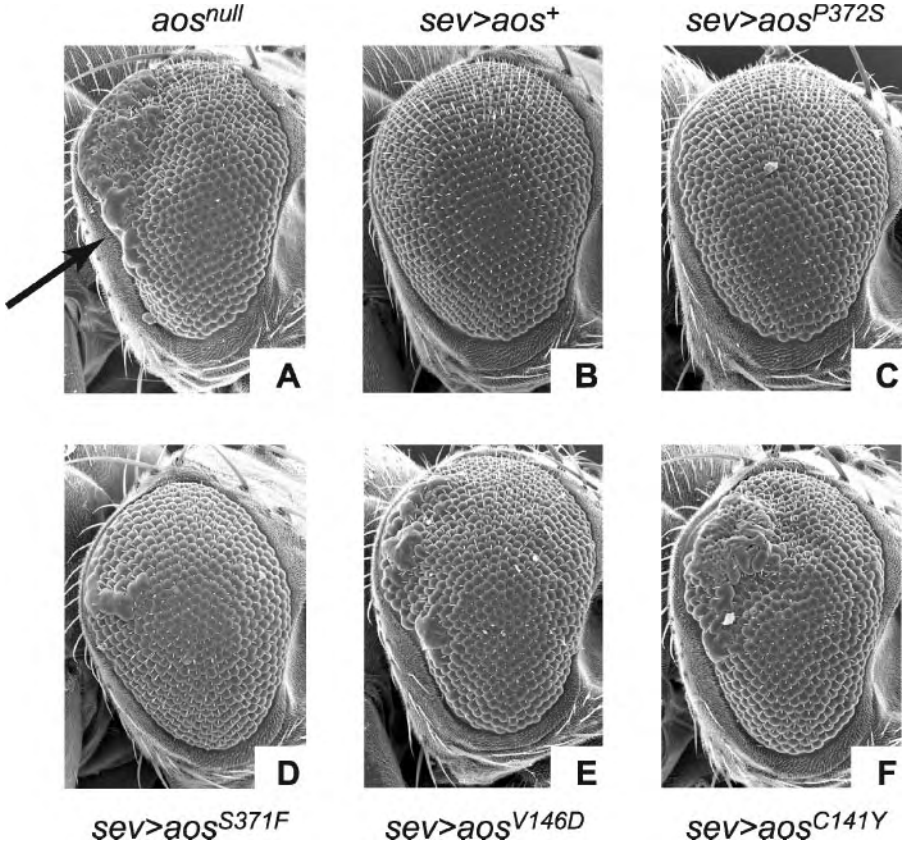


FIGURE 5. **Misexpression of Aos mutants differentially rescue loss of endogenous *aos*.** Scanning electron micrographs of adult compound eyes lacking endogenous *aos* activity (*A*), with a transgene encoding the indicated *aos* allele and the *sev-GAL4* driver (*B–F*). The posterior margin of the eye is marked in *A* by a *black arrow*. *A* is a scanning electron micrograph of an eye lacking all endogenous *aos* activity (*aos$^{null}$*; *aos$^{\Delta 7}$/Df(3L)Exel6129*), whereas *B* demonstrates an almost complete restoration of the disrupted eye morphology to wild type by an *aos* transgene driven by the *sev-GAL4* driver (*B*, *aos$^{null}$ sev>aos$^{+}$*). In *C* (*aos$^{null}$* sev>*aos$^{P372S}$*), the morphology of the *aos$^{null}$* eye was also greatly improved, suggesting that this class I allele retains substantial rescuing capability. The morphology of *aos$^{null}$* eyes expressing the remaining class I alleles (*D*, aos$^{null}$ sev>*aos$^{S371F}$*; *E*, *aos$^{null}$* sev>*aos$^{V146D}$*) was only moderately improved, suggesting they both retain a measure of activity, with *aos$^{S371F}$* consistently displaying slightly better morphology. In contrast, the morphology of *aos$^{null}$* eyes was not improved by expression of the class II allele *aos$^{C141Y}$*, consistent with a lack of activity (*F*, *aos$^{null}$* sev>*aos$^{C141Y}$*).

mutated *aos* transgenes showed similar activity profiles in other tissues (Table 1), indicating that the suppressor lines from which they were derived did not have tissue-specific alterations. The one exception to this is *aos$^{P372S}$*, which showed only partial suppression in the eye (and in viability) but apparently strong suppression in the wing (Table 1).

*Molecular Characterization of aos Alleles*—To ascertain the molecular nature of each mutation, we sequenced the *aos* transgene in each of the 16 suppressor lines (Table 1). We grouped the sequenced alleles into three general classes, based on the nature of the mutation uncovered (Table 1 and Fig. 2*B*). We recovered three missense mutations in residues other than cysteine, which we named class I alleles. We designated as class II alleles missense mutations that alter conserved cysteines and as class III alleles nonsense mutations that generated truncated gene products. To avoid confusion, all of the alleles will be henceforth referred to by their amino acid change. All three allelic classes highlight the relative importance of both the NCR and the CCR. Class III alleles delete all or a portion of the CCR and lack activity in our misexpression assays (Figs. 2*B* and 3*G* and Table 1), in agreement with the findings with C-terminal truncations engineered by Howes *et al.* (26). Class II alleles (cysteine mutants) behaved as nearly complete suppressors, consistent with the notion that proper Aos folding depends upon the formation of a correct pattern of disulfide bonds, including those outside the putative EGF domain. A cysteine in the NCR (C141) was altered in two alleles, and three cysteines in the CCR (C341, C407, and C413) were mutated (Figs. 2*B* and 3*F* and Table 1). Class I mutants functionally implicate both the NCR and CCR in Aos activity. *aos$^{S371F}$* and *aos$^{P372S}$* affect well conserved neighboring residues in the CCR but display different strengths *in vivo* (Fig. 2*B* and 3, *C* and *D*, and Table 1). *aos$^{P372S}$* retains moderate activity in the eye and embryo (the latter inferred from viability studies), whereas *aos$^{S371F}$* exhibits minimal activity in the misexpression assays. The only class I allele uncovered in the NCR, *aos$^{V146D}$*, alters a conserved valine (isoleucine in TcAos) to aspartate and displays minimal activity in the misexpression assays (Figs. 2*B* and 3*E* and Table 1). Western blot analysis confirmed that the differences in allelic activity are due neither to nonsense-mediated decay nor decreased protein stability. Misexpression of *aos* mutants in the follicle cells of the ovary (using a *CY2-*

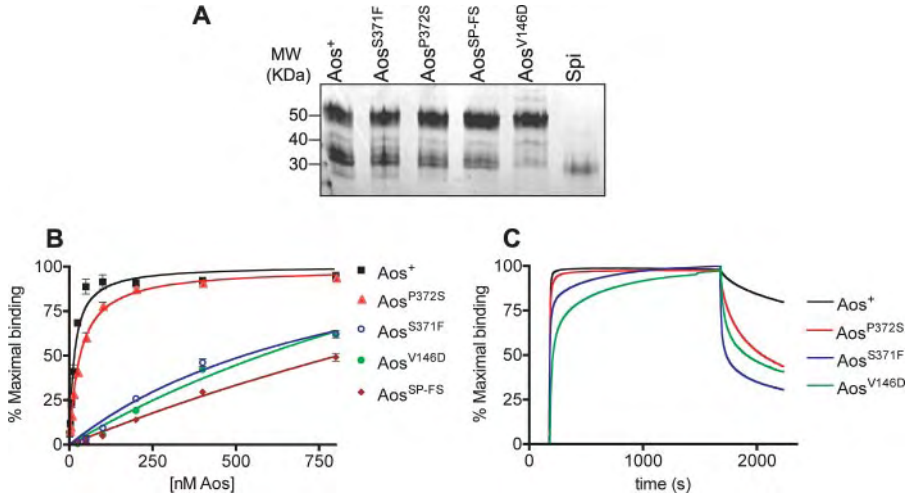


FIGURE 6. **Aos mutants display compromised binding to Spi.** *A*, Coomassie-stained SDS-PAGE gel of purified proteins utilized in these studies. Proteins were loaded at a concentration of ~0.25 mg/ml. A degradation product corresponding to the C terminus of Aos was observed in all Aos variants, as determined by immunoblotting against the C-terminal His tag (not shown). *B*, SPR experiments reveal decreased Spi binding affinities by Aos mutants. Increasing concentrations of Aos mutants (analytes) were flowed over sensorchips with immobilized Spi (ligand), and the normalized responses at equilibrium ($B_{max}$) were plotted as a function of analyte concentration in nM. $Aos^{P372S}$ binds strongly to Spi with a $K_D$ of ~32 nM, only 2-fold weaker than $Aos^{+}$. $Aos^{S371F}$ and $Aos^{V146D}$ bind Spi much more weakly, with $K_D$ values of ~555 and ~646 nM, respectively. The double mutant $Aos^{SP-FS}$ bound Spi least strongly, with a $K_D$ of ~834 nM. *C*, SPR sensorgrams show a qualitative difference in the on and off binding rates of Spi to $Aos^{+}$ *versus* Aos mutants. Aos proteins were injected at 800 nM onto a Spi sensorchip, and responses were normalized against their respective maximal binding and plotted as a function of time. A clear increase in the off rate can be observed with $Aos^{P372S}$, which binds ~2-fold more weakly to Spi than $Aos^{+}$.

**TABLE 2**
**Summary of mutant Aos affinities for Spi**

| Allele | Spi $K_D$ | Fold decrease in $K_D$ over wild type |
|---|---|---|
| | *nM* | |
| $Aos^{+}$ | 15.0 ± 1.4 | 1 |
| $Aos^{P372S}$ (*aos*$^{334}$) | 32.3 ± 0.9 | 2.1 |
| $Aos^{S371F}$ (*aos*$^{39}$) | 554 ± 48 | 36.9 |
| $Aos^{V146D}$ (*aos*$^{136}$) | 646 ± 88 | 43.1 |
| $Aos^{SP-FS}$ | 834 ± 92 | 55.6 |

*GAL4* driver), where Aos has also been reported to inhibit dEGFR signaling (27), led to high levels of protein expression (Fig. 4). The sole exception to this was *aos*$^{Q107X}$, which encodes an early nonsense mutation that eliminates the Aos antibody epitope (Fig. 4).

As a more accurate and sensitive indicator of the physiological activity of the class I mutations relative to wild type Aos, we also assessed their ability to compensate for the loss of endogenous Aos in the eye. Loss of *aos* activity in the eye results in a roughened appearance along with a prominent blistering along the posterior margin of the adult compound eye (Fig. 5*A*) (29). Misexpression of wild type *aos* in developing photoreceptors (using *sevenless-GAL4*) in an *aos* mutant background rescues the knock-out phenotype and restores the eye to nearly wild type morphology (Fig. 5*B*) (26, 29). Consistent with the observation that *aos*$^{P372S}$ retains considerable activity in the suppression assays, we observed significant (but not complete) rescue of the *aos* null phenotype with this allele (Fig. 5*C*). We also noted differences in the relative abilities of the other mutations to rescue the *aos* null eye phenotype. Both *aos*$^{S371F}$ and *aos*$^{V146D}$ provide weak rescue, with *aos*$^{S371F}$ appearing to have slightly more activity, indicating that these alleles retain some degree of activity. In contrast, the class II mutant *aos*$^{C141Y}$, which behaved as a complete suppressor in the screen, was unable to rescue the *aos* null phenotype in the eye, consistent with its lack of activity (and likely misfolding).

*Aos Mutants Are Defective in Their Ability to Bind and Inhibit Spi*—Given the recently identified action of Aos as a ligand sink, we next asked whether the loss of activity displayed by the Aos alleles was due to a reduced affinity for Spi. To address this question, we purified recombinant Aos proteins encoding the class I mutations uncovered in the screen (Fig. 6*A*) and assessed their ability to bind immobilized Spi using SPR, as described under "Experimental Procedures." We excluded from this analysis mutants that modified cysteines or incorporated premature stop codons, because these alleles were inactive *in vivo* and are likely to produce proteins with folding defects that would be difficult to purify. We found that each class I Aos allele exhibited a weaker Spi binding affinity than the wild type protein, showing an overall correspondence with its phenotypic strength (Fig. 6*B* and Table 2). The strong suppressor $Aos^{V146D}$, which maps to the NCR, bound Spi with an apparent $K_D$ of 646 nM (~43-fold weaker than wild type), suggesting that the NCR contributes substantially to Aos function both *in vitro* and *in vivo*. In the CCR, the strong suppressor $Aos^{S371F}$ bound with an apparent $K_D$ of 554 nM (~37-fold weaker than wild type). In contrast, $Aos^{P372S}$, which behaved as a moderate suppressor *in vivo*, bound Spi with an apparent $K_D$ of 32 nM, indicating that just a 2-fold decrease in affinity is sufficient to interfere with Aos function *in vivo*. This small decrease in affinity correlated with an increased off rate in SPR sensorgrams (Fig. 6*C*). We also generated a double mutant encoding both class I changes in the CCR ($Aos^{SP-FS}$) and found that the effects of the single mutants on ΔG for Spi binding are close to additive, because $Aos^{SP-FS}$ bound Spi with a $K_D$ of 834 nM (about 55-fold weaker than wild type).

Finally, we investigated the relative abilities of the Aos mutants to inhibit Spi-dependent dEGFR phosphorylation in cultured cells. We treated S2 cells expressing low levels of dEGFR (to prevent self-activation) with 50 nM Spi, titrated in increasing amounts of purified Aos proteins (wild type or mutated), and examined dEGFR tyrosine phosphorylation levels by Western blotting (Fig. 7). In agreement with our studies *in vivo* and by SPR, the stronger alleles ($Aos^{S371F}$ and $Aos^{V146D}$) were unable to inhibit Spi-dependent dEGFR phosphorylation, even when used at a 25-fold molar excess over Spi. In contrast,

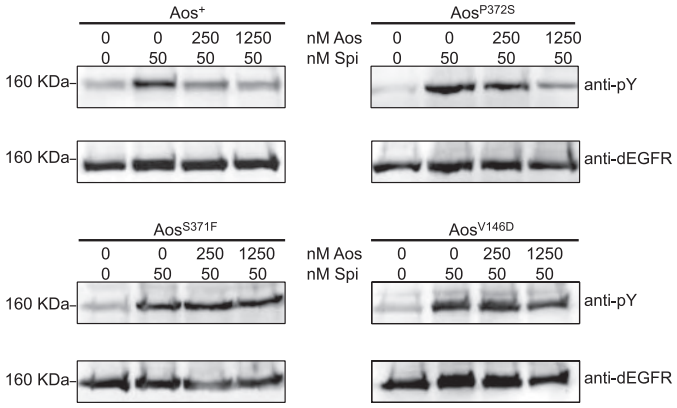


FIGURE 7. **Aos mutants are deficient in their ability to inhibit Spi-dependent dEGFR phosphorylation.** D2f cells expressing low levels of dEGFR were left unstimulated or were stimulated on ice for 10 min with 50 nM Spi together with increasing concentrations of purified wild type or mutated Aos proteins. dEGFR tyrosine phosphorylation levels were assayed by immunoblotting with anti-phosphotyrosine (*top panels*). Immunoblots were stripped and reprobed with anti-dEGFR to ensure equal loading (*bottom panels*). Whereas a 5-fold molar excess of $Aos^{+}$ over Spi is sufficient to abolish dEGFR phosphorylation, a 25-fold excess of $Aos^{P372S}$ is necessary to decrease dEGFR phosphorylation to nearly base-line levels. In contrast, $Aos^{S371F}$ and $Aos^{V146D}$ do not affect Spi activity even at high molar excesses.

$Aos^{P372S}$ nearly abolished dEGFR phosphorylation when added at a 25-fold excess, compared with the 5-fold excess of wild type Aos required to completely abolish dEGFR phosphorylation (Fig. 7) (25). These results demonstrate that the loss of activity of Aos mutants is due to decreased binding affinities for Spi.

## DISCUSSION

We have identified a series of point mutations in Aos that exhibit a range of phenotypes and have effects in multiple developmental contexts. We have also correlated the phenotypic strength of these mutants with decreased *in vitro* Spi binding affinity of the encoded Aos protein. The point mutations uncovered in this study implicate both the N- and C-terminal cysteine-rich regions of Aos in Spi binding and appear to define an affinity threshold for Aos function during development.

*dEGFR Signaling Is Highly Susceptible to Small Changes in the Affinity of Aos for Spi*—Recent computational studies by Reeves *et al.* (32) suggested that Aos confers robustness to the dEGFR signaling module in *Drosophila* development by restricting the range of Spi action and predicted that highly efficient sequestration of Spi by Aos is necessary to generate a robust feedback loop. It follows that the system should be highly susceptible to small changes in the Aos-Spi affinity or dissociation rate, and our data are in agreement with this prediction. Remarkably, a decrease of just ~2-fold in the affinity of Aos for Spi (caused by $Aos^{P372S}$) as assessed in our *in vitro* studies is sufficient to reduce the *in vivo* activity of this mutant in the developing eye and embryo and to greatly diminish it in the adult wing. Decreases of 37- or 43-fold in Spi binding affinity, seen with $Aos^{S371F}$ and $Aos^{V146D}$, respectively, abolished most Aos activity from all of the tested tissues.

We also addressed whether the Aos mutants could rescue an *aos* null phenotype, to control for potential genetic biases inherent in the suppression studies. The rescue data support the notion that a modest decrease in binding affinity for Spi is sufficient to compromise Aos activity *in vivo*, because Aos mutants exhibited a similar (although overall higher) pattern of activity than in the suppression assay. The higher level of Aos mutant activity in the rescue assay may reflect a higher sensitivity of the *aos* null genetic background to Aos protein levels. Importantly, the progressive loss of activity of Aos alleles seen in flies could also be recapitulated in a cultured cell system. In agreement with genetic data, inhibition of Spi-induced dEGFR phosphorylation required a larger excess of purified $Aos^{P372S}$ than wild type Aos. In the cases of $Aos^{S371F}$ and $Aos^{V146D}$, we were unable to observe any inhibition even when added at a 25-fold excess over Spi. Our results argue that the high binding affinity between Aos and Spi (and perhaps other ligands) is a critical parameter for full Aos activity *in vivo* and for conferring dEGFR signaling robustness. It will be interesting to determine whether dEGFR signaling is similarly susceptible to small changes in the affinity of dEGFR for its ligands.

One Aos mutant ($Aos^{P372S}$) exhibited a more complete suppression phenotype in the wing than in the eye or in viability studies. This could reflect the strength of the GAL4 driver utilized in the wing (*MS1096*). Another more intriguing possibility is that Vein, which appears to be the primary dEGFR ligand in the wing (where Spi plays no role (41)) is inhibited less efficiently by Aos, and Aos mutations would therefore impair its effect more completely. However, the effects of Aos on other dEGFR ligands have not yet been reported.

*Structural Considerations*—Aos was originally thought to inhibit dEGFR directly, via a proposed "atypical" EGF domain. It is now known instead to act as a ligand antagonist, sequestering Spi in a manner similar to bone morphogenetic protein inhibition by Chordin (orthologous to *Drosophila* Short gastrulation) (42–45), and insulin growth factor 1 binding by the insulin growth factor-binding protein family (46). A common feature of these ligand antagonists is the presence of two cysteine-rich domains. Our results suggest that the two cysteine-rich domains in Aos, NCR and CCR, function interdependently in binding and sequestering Spi. Indeed, mutations in either the NCR or CCR impair Aos function and *in vitro* Spi binding. Interestingly, $Aos^{V146D}$ binds Spi with a $K_D$ similar to that previously reported for an Aos fragment containing only the C-terminal 225 amino acids (and lacking the NCR altogether) (25). This observation suggests that the C-terminal region of Aos provides the majority of Spi binding interactions, with additional weak contributions from the NCR being required for high affinity binding of full-length Aos to Spi. A similar bipartite binding site involving two cysteine-rich regions has been reported for insulin growth factor-binding protein family members (47). It is interesting to speculate that the required contributions of both the NCR and CCR for efficient Spi sequestration, and their linkage by a proteolytically sensitive region could support a proteolysis-based mechanism for relieving Spi inhibition (or attenuating Aos function) during development. A precedent for this exists in bone morphogenetic protein signaling where the protease Tolloid relieves Chordin-mediated sequestration of bone morphogenetic protein (48).

The findings in this report show a clear correlation between the failure of Aos to function *in vivo* and a reduction in its ability to efficiently sequester Spi *in vitro* and in cell culture systems. Thus, our studies provide direct *in vivo* evi-

dence for the Spi sequestration model presented by Klein *et al.* (25) based on *in vitro* analyses of interactions between Aos, Spi, and the dEGFR extracellular region. In addition, the fact that missense mutations in both the N- and C-terminal cysteine-rich regions of Aos impair Spi binding suggests a bipartite Spi-binding site in Aos and provides the first insight into how this ligand antagonist achieves its function. A full understanding of how Aos recognizes its target and how this can be exploited in the design of antagonists for human EGF ligands will require crystallographic and further functional studies of the Aos-Spi complex.

*Acknowledgments—We thank Rudi Turner for capturing scanning electron microscopy images and Stas Shvartsman, Gregory Reeves, Kim Cook, Daryl Klein, Jeannine Mendrola, and other members of the Lemmon and Duffy laboratories for constructive advice with experiments and the manuscript. We also thank the Developmental Studies Hybridoma Bank for antibodies and the Bloomington Stock Center for fly stocks.*

## REFERENCES

1. Ben-Shlomo, I., Yu Hsu, S., Rauch, R., Kowalski, H. W., and Hsueh, A. J. (2003) *Sci. STKE 2003*, RE9
2. Stein, R. A., and Staros, J. V. (2000) *J. Mol. Evol.* **50,** 397–412
3. Arteaga, C. L. (2003) *Exp. Cell Res.* **284,** 122–130
4. Graham, J., Muhsin, M., and Kirkpatrick, P. (2004) *Nat. Rev. Drug Discov.* **3,** 549–550
5. Holbro, T., Civenni, G., and Hynes, N. E. (2003) *Exp. Cell Res.* **284,** 99–110
6. Cohen, M. H., Williams, G. A., Sridhara, R., Chen, G., McGuinn, W. D., Jr., Morse, D., Abraham, S., Rahman, A., Liang, C., Lostritto, R., Baird, A., and Pazdur, R. (2004) *Clin. Cancer Res.* **10,** 1212–1218
7. Gazdar, A. F., Shigematsu, H., Herz, J., and Minna, J. D. (2004) *Trends Mol. Med.* **10,** 481–486
8. Shilo, B. Z. (2005) *Development* **132,** 4017–4027
9. Gonzalez-Reyes, A., and St. Johnston, D. (1998) *Development* **125,** 2837–2846
10. Nilson, L. A., and Schupbach, T. (1999) *Curr. Top Dev. Biol.* **44,** 203–243
11. Chang, J., Jeon, S. H., and Kim, S. H. (2003) *Mol. Cell* **15,** 186–193
12. Chang, J., Kim, I. O., Ahn, J. S., and Kim, S. H. (2001) *Int. J. Dev. Biol.* **45,** 715–724
13. Golembo, M., Raz, E., and Shilo, B. Z. (1996) *Development* **122,** 3363–3370
14. Golembo, M., Yarnitzky, T., Volk, T., and Shilo, B. Z. (1999) *Genes Dev.* **13,** 158–162
15. Guichard, A., Biehs, B., Sturtevant, M. A., Wickline, L., Chacko, J., Howard, K., and Bier, E. (1999) *Development* **126,** 2663–2676
16. Simcox, A. A., Grumbling, G., Schnepp, B., Bennington-Mathias, C., Hersperger, E., and Shearn, A. (1996) *Dev. Biol.* **177,** 475–489
17. Wessells, R. J., Grumbling, G., Donaldson, T., Wang, S. H., and Simcox, A. (1999) *Dev. Biol.* **216,** 243–259
18. Tio, M., Ma, C., and Moses, K. (1994) *Mech. Dev.* **48,** 13–23
19. Tio, M., and Moses, K. (1997) *Development* **124,** 343–351
20. Martin-Blanco, E., Roch, F., Noll, E., Baonza, A., Duffy, J. B., and Perrimon, N. (1999) *Development* **126,** 5739–5747
21. Reich, A., and Shilo, B. Z. (2002) *EMBO J.* **21,** 4287–4296
22. Ghiglione, C., Carraway, K. L., 3rd, Amundadottir, L. T., Boswell, R. E., Perrimon, N., and Duffy, J. B. (1999) *Cell* **96,** 847–856
23. Alvarado, D., Rice, A. H., and Duffy, J. B. (2004) *Genetics* **167,** 187–202
24. Freeman, M., Klambt, C., Goodman, C. S., and Rubin, G. M. (1992) *Cell* **69,** 963–975
25. Klein, D. E., Nappi, V. M., Reeves, G. T., Shvartsman, S. Y., and Lemmon, M. A. (2004) *Nature* **430,** 1040–1044
26. Howes, R., Wasserman, J. D., and Freeman, M. (1998) *J. Biol. Chem.* **273,** 4275–4281
27. Wasserman, J. D., and Freeman, M. (1998) *Cell* **95,** 355–364
28. Golembo, M., Schweitzer, R., Freeman, M., and Shilo, B. Z. (1996) *Development* **122,** 223–230
29. Freeman, M. (1994) *Development* **120,** 2297–2304
30. Brodu, V., Elstob, P. R., and Gould, A. P. (2004) *Dev. Cell* **7,** 885–895
31. Freeman, M. (1997) *Development* **124,** 261–270
32. Reeves, G. T., Kalifa, R., Klein, D. E., Lemmon, M. A., and Shvartsman, S. Y. (2005) *Dev. Biol.* **284,** 523–535
33. Ashburner, M. (1989) *Drosophila: A Laboratory Handbook*, Cold Spring Harbor, New York
34. Schweitzer, R., Howes, R., Smith, R., Shilo, B. Z., and Freeman, M. (1995) *Nature* **376,** 699–702
35. MacLaren, C. M., Evans, T. A., Alvarado, D., and Duffy, J. B. (2004) *Dev. Genes Evol.* **214,** 360–366
36. Gur, G., Rubin, C., Katz, M., Amit, I., Citri, A., Nilsson, J., Amariglio, N., Henriksson, R., Rechavi, G., Hedman, H., Wides, R., and Yarden, Y. (2004) *EMBO J.* **23,** 3270–3281
37. Brand, A. H., and Perrimon, N. (1993) *Development* **118,** 401–415
38. Casci, T., Vinos, J., and Freeman, M. (1999) *Cell* **96,** 655–665
39. Taguchi, A., Sawamoto, K., and Okano, H. (2000) *Genetics* **154,** 1639–1648
40. Duffy, J. B. (2002) *Genesis* **34,** 1–15
41. Simcox, A. (1997) *Mech. Dev.* **62,** 41–50
42. Garcia Abreu, J., Coffinier, C., Larrain, J., Oelgeschlager, M., and De Robertis, E. M. (2002) *Gene* (*Amst.*) **287,** 39–47
43. Francois, V., and Bier, E. (1995) *Cell* **80,** 19–20
44. Sasai, Y., Lu, B., Steinbeisser, H., and De Robertis, E. M. (1995) *Nature* **376,** 333–336
45. Piccolo, S., Sasai, Y., Lu, B., and De Robertis, E. M. (1996) *Cell* **86,** 589–598
46. Clemmons, D. R. (2001) *Endocr. Rev.* **22,** 800–817
47. Siwanowicz, I., Popowicz, G. M., Wisniewska, M., Huber, R., Kuenkele, K. P., Lang, K., Engh, R. A., and Holak, T. A. (2005) *Structure* **13,** 155–167
48. Piccolo, S., Agius, E., Lu, B., Goodman, S., Dale, L., and De Robertis, E. M. (1997) *Cell* **91,** 407–416

***Nature Manuscript 2008-01-00681: in press***

# Structural basis for EGFR ligand sequestration by Argos

Daryl E. Klein, Steven E. Stayrook, Fumin Shi, Kartik Narayan, and Mark A. Lemmon

*Department of Biochemistry and Biophysics, University of Pennsylvania School of Medicine, 809C Stellar-Chance Laboratories, 422 Curie Boulevard, Philadelphia, PA 19104-6059, U.S.A.*

**Members of the Epidermal Growth Factor Receptor (EGFR) or ErbB/HER family and their activating ligands are essential regulators of diverse developmental processes[1,2]. Inappropriate activation of these receptors is a key feature of many human cancers[3], and its reversal is an important clinical goal. A natural secreted antagonist of EGFR signalling, called Argos, was identified in *Drosophila*[4]. We showed previously that Argos functions by directly binding (and sequestering) growth factor ligands that activate EGFR[5]. Here we describe the 1.6Å resolution crystal structure of Argos bound to an EGFR ligand. Contrary to expectations[4,6], Argos contains no EGF-like domain. Instead, a trio of closely-related domains (which resemble a three-finger toxin fold[7]) form a clamp-like structure around the bound EGF ligand. Although structurally unrelated to the receptor, Argos mimics EGFR by using a bipartite binding surface to entrap EGF. The individual Argos domains share unexpected structural similarities with the extracellular ligand-binding regions of TGFβ family receptors[8]. The three-domain clamp of Argos also resembles the urokinase-type plasminogen activator (uPA) receptor, which uses a similar mechanism to engulf the uPA EGF-like module[9]. Our results suggest that undiscovered mammalian counterparts of Argos may exist among other poorly**

**characterized structural homologues. In addition, the structures presented here define requirements for the design of artificial EGF-sequestering proteins that would be valuable anti-cancer therapeutics.**

The 419-residue mature *D. melanogaster* Argos was modified to improve crystallization by removing a poorly conserved region with multiple O-linked glycosylation sites (residues 1-87). In addition, a non-conserved 120-amino acid insert (residues 140-259) exclusive to drosophilid Argos molecules was replaced with the corresponding 5 amino acids (PDGRT) from *Apis mellifera* Argos (Supplementary Fig. 1). The resulting 217-amino acid protein ($Argos_{217}$) exhibits greatly improved stability and purification properties, and binds Spitz with an affinity ($K_D$ = 7.7nM: Supplementary Fig. 2) similar to that previously measured[5] for full-length $Argos_{419}$ ($K_D$ = 20nM). Triclinic crystals of $Argos_{217}$ bound to the Spitz EGF-domain ($Spitz_{EGF}$: residues 48-99) grew at neutral pH with two complexes per asymmetric unit, and diffracted to 1.6Å resolution. The structure was solved by multiwavelength anomalous dispersion (MAD), using the halide soak method[10] (Supplementary Table 1). Representative electron density is shown in Fig. 1a,b. Structures of uncomplexed $Argos_{217}$ and $Spitz_{EGF}$ were also determined (to 2.5Å and 1.5Å respectively) by molecular replacement.

Argos consists of three separate disulphide-bonded β-sheet domains (domains 1-3) that show no resemblance to EGF domains. This three-domain composition was not discerned in sequence analyses. The three domains of Argos engulf the bound $Spitz_{EGF}$ molecule with a structure that is reminiscent of a C-clamp (Fig. 1c,d). Domains 2 and 3 constitute the 'jaws' of this clamp, and make an intimate set of direct contacts with bound $Spitz_{EGF}$. Domain 1 forms the backbone of the C-clamp and does not contribute directly to ligand binding. $Spitz_{EGF}$ itself is very similar in structure to other known EGFR family

ligands (Supplementary Fig. 3). Its three disulphide bonds generate three loops in the typical EGF domain structure, which are termed the A-, B-, and C-loops (Fig. 1e). The Spitz B-loop protrudes into the crevice between domains 2 and 3 of Argos (Fig. 1c,d). The conformation of $Spitz_{EGF}$ is largely unaltered upon binding to Argos, apart from small changes in the backbone at the B-loop tip (Supplementary Fig. 3) and reorientation of certain interfacial side-chains.

Domains 1, 2 and 3 of Argos superimpose remarkably well (with $C_\alpha$ r.m.s. deviations of 1.3-1.9Å) despite sharing little sequence identity (<30%). The overall architecture found in all 3 domains is shown in Fig. 2a. The cysteines form a disulphide-bonded core, from which two β-hairpins project to form a four-stranded β-sheet with a relatively unusual antiparallel 2143 topology[11]. The result is a flat domain that resembles part of a left hand, with the β-hairpins as two fingers plus a thumb-like projection emerging from the disulphide-bonded core. In domains 1 and 3 (but not domain 2), a knuckle-like protrusion also projects below the plane of the page in Fig. 2a. The positions of the C1-C3 and C4-C6 disulphides are almost identical in all three Argos domains (Fig. 2b), but the third (C2-C5) disulphide is missing from domain 2. The absence of this disulphide correlates with the lack of a knuckle in domain 2. Domain 1 is distinguished by the presence of a unique additional N-terminal β-strand (β1`) that is parallel to strand β2. The drosophilid-specific insertion in Argos occurs at the top of domain 1 in the orientation shown in Fig. 2, immediately before the knuckle (between C4 and C5). This insertion would likely project out and away from domain 1, with its ends constrained by the C2-C5 and C4-C6 disulphides.

Protein Data Bank searches with the DALI server[12] and a secondary structure matching (SSM) algorithm[13] revealed that the three domains of Argos are significantly

related to the three-finger toxin fold found in snake neurotoxins and cardiotoxins[7] – although the disulphide-bonding pattern is altered, and Argos has just two (rather than three) fingers per domain. Interestingly, the three-finger toxin fold is also found in the extracellular ligand-binding domains of receptors for TGFβ family ligands[8]. As shown in Fig. 2c, two fingers of the extracellular ligand-binding domain of the type II activin receptor (ActRII)[14] overlay very well with domain 3 of Argos. Members of the Ly-6 superfamily also share this fold[7], including the receptor for urokinase-type plasminogen activator (uPA)[9].

Argos 'clamps' $Spitz_{EGF}$ between domains 2 and 3 (Fig. 1c,d), and buries 35% of the ligand's surface. Domains 2 and 3 are approximately parallel to one another (Fig. 1c), and are stacked so that they present opposite surfaces to the $Spitz_{EGF}$ molecule sandwiched between them. Using the hand analogy introduced above, domain 3 presents its palm to the bound ligand, whereas domain 2 contacts $Spitz_{EGF}$ using the back of the hand. The Spitz-binding regions on domains 2 and 3 can each be divided into two sites (Fig. 3): an A-site and a B-site. The A sites (d2A, d3A) lie on the β-sheet surfaces, and the B sites (d2B, d3B) involve the thumb and disulphide-bonded core. Site A on domain 2 (d2A) consists of a patch of hydrophobic side-chains on the back-of-hand surface of fingers 1 and 2 (L301, L326, Y341, F343), which makes van der Waal's contacts with three aliphatic side-chains from the Spitz B-loop (V72, I74 and V79). Site B on domain 2 (d2B) comprises a flat surface (cyan in Fig. 3) formed by side-chains from the base of finger 2 (V323, Y325, S346 and P347) and the domain 2 thumb (F294). This flat surface packs against the C-terminal part of $Spitz_{EGF}$, contacting M89, Q91, and Y95 in the C-loop. Importantly, we previously identified S346 and P347 from site d2B as important residues in a genetic screen for modifiers of an Argos misexpression phenotype in Drosophila eye development[15]. S346 and P347 are both conserved in all known Argos orthologs (Supplementary Fig. 1), and

make direct hydrogen bonds with Spitz (Fig. 3). In domain 3, part of site A (d3A) involves polar side-chains on the palm side of finger 1 (T363, R365, E373, N375) that interact with the $Spitz_{EGF}$ B-loop. In addition, a cluster of hydrophobic side-chains around the tip of finger 2 (F403, L404, I408) contacts residues from the $Spitz_{EGF}$ N-terminus, A-loop, and B-loop. A key feature of this interaction is the projection of the F403 and L404 side-chains into a hydrophobic pocket on the Spitz surface formed by Y52, P55, P78, and Y80 (Fig. 3). In the B-site on domain 3 (d3B), side-chains close to the base of the two fingers (Q357, P358, L361, N377, and S412) form a binding site for the Spitz A-loop helix (contacting Spitz T57, F58, W61, and Y62). Site d3B also accommodates the side-chains of Spitz R92 and L64 (Fig. 3).

Despite having a completely different structural scaffold from EGFR, Argos mimics the characteristic bipartite capture of growth factor seen in ligand-bound structures of the EGFR extracellular region (sEGFR)[16,17]. Argos presents two ligand-binding surfaces that closely resemble those in EGFR. Specifically, domain 2 of Argos mimics domain I of sEGFR in its ligand contacts, whereas domain 3 of Argos emulates sEGFR domain III (Fig. 4a). The primary ligand contacts made by domain 2 of Argos and domain I of EGFR are remarkably similar. Both utilize a central hydrophobic patch that interacts with a similar region of the B-loop of the bound EGF domain. As shown in Fig. 4a (upper panels), Argos site d2A and EGFR site 1 (from Ogiso *et al.*[17]) are very similar – with a comparable arrangement of hydrophobic side-chains making ligand contacts in each case. The second binding site on Argos domain 2 (d2B) also recapitulates many other sEGFR domain I/hEGF contacts, but is different in detail. Along the same lines, site B on Argos domain 3 (d3B) recapitulates site 2 in the sEGFR/hEGF interface[17] (Fig. 4a, lower panels), including interactions with an arginine that is critical for hEGF binding to its receptor (R41 in hEGF, R92 in Spitz). Both sites accommodate three key ligand side-chains (Y13, L15 and R41 in

hEGFR: Y62, L64 and R92 in Spitz) in analogous binding sites. Site 3 on domain III of EGFR is not mimicked by Argos; the C-terminus of the bound EGF domain is much more exposed in the $Spitz_{EGF}$/Argos complex than in the hEGF/sEGFR complex (Fig. 4a). Argos compensates for the absence of these site 3 interactions with an extensive set of unique contacts mediated by site d3A (Fig. 4a).

Overall, Argos domain 2 buries slightly less surface on $Spitz_{EGF}$ (560Å$^2$) than EGFR domain I buries on hEGF (745Å$^2$)[17], but includes more apolar surface (68%) than in the EGFR domain/hEGF interface (56%). Argos domain 3 buries slightly more of its bound EGF domain (843Å$^2$: 65% apolar) than does sEGFR domain III (819Å$^2$: 62% apolar). Each interface has a shape complementarity parameter ($S_c$)[18] of 0.70, which is typical for strong protein/protein interfaces and reflects a significantly greater shape complementarity than seen in antibody/antigen interfaces. The fact that $Spitz_{EGF}$ binds ~10-fold more strongly to Argos than to the Drosophila EGFR extracellular region may reflect different requirements for domain rearrangements in the two binding proteins. Clearly defined sets of (intramolecular) domain1/2 and domain 1/3 contacts (Supplementary Fig. 4) may optimize bipartite capture of $Spitz_{EGF}$ (or inhibit $Spitz_{EGF}$ dissociation). A crystal structure of unliganded $Argos_{217}$ (Supplementary Figs. 5 & 6) suggests that domain 3 may be mobile, possibly ‘collapsing’ onto domain 2 in the absence of bound ligand (see Supplementary Information for discussion). Once $Spitz_{EGF}$ is bound, domain 1/3 interactions in the complex may slow down its dissociation. Interestingly, approximately half of the residues involved in $Spitz_{EGF}$ binding to Argos are conserved in human EGF and/or TGFα, which may explain the ability of Argos to bind detectably (although weakly, with $K_D$~5μM) to human EGF (not shown).

Sequence analyses have failed to identify clear homologues of Argos in vertebrates, but this does not necessarily mean that functional mammalian analogues do not exist. The amino acid sequence of Argos has been unusually cryptic, providing few (or misleading) clues about the structure of the protein. It only became apparent that Argos has three distinct domains (and no EGF-like domain) once the structure described here was determined. Moreover, the relationship of the constituent domains in Argos to the three-finger toxin fold can only been seen in structural (and not sequence) comparisons. As shown in Fig. 2c, the individual domains of Argos share unexpected and striking structural similarity with the extracellular ligand-binding regions of receptors for TGFβ/bone morphogenetic protein (BMP) family ligands, which consist of little more than a single three-finger toxin fold[8,14]. As shown in Supplementary Fig. 7, the positions of the ligand-binding sites in Argos domain 3 and the extracellular regions of TGFβ/BMP family receptors also correspond strikingly well. Both utilize the palm side of the domain according to the analogy drawn in Fig. 2a.

Among human proteins that contain three-finger toxin fold domains, one intriguing example uses three such modules to engulf an EGF-like domain. This is the cell-surface receptor for urokinase plasminogen activator (uPA)[9,19], which represents an unexpected structural homologue of Argos. As shown in Fig. 4b, the three domains from the uPA receptor (uPAR) form a clamp-like structure around the EGF domain found at the amino terminus of uPA. Each domain presents its palm side to the bound ligand (like Argos domain 3). The uPAR structure resembles the clamp formed around the Spitz EGF domain by Argos (Fig. 1c,d). There are differences in the orientation of the bound EGF domain in the Argos/Spitz and uPAR/uPA complexes. Moreover, Argos has one of its three constituent domains (domain 2) ‘inverted’ so that it presents the back (rather than palm) of the hand to the ligand. However, the correspondence in overall architecture and function

(as proteins that entrap EGF domains) of uPAR and Argos suggest that other structural homologues of Argos should be sought in mammals. There are many human uPAR/Ly6 domain-containing proteins for which the function remains unclear. Several, such as CD177/PRV-1 and C4.4A, contain multiple three-finger domains[20,21] like uPAR. Moreover, C4.4A expression is known to be altered in several metastatic human cancers[22]. We suggest that one of these numerous structural homologues might represent a functional analogue of Argos. Even if such an analogue does not exist, the known human proteins from this class could clearly be used as structural scaffolds in the design of protein therapeutics that will sequester ErbB receptor-activating EGF domains.

It is increasingly clear that excessive or unregulated expression (or shedding) of ErbB family ligands is important in numerous cancers, through autocrine and/or paracrine activation of cell growth[23-27]. The role of ErbB ligands may be particularly important in cancers where available receptor-targeted approaches have failed or have met resistance[26]. In these (and other) cases, therapeutic agents that neutralize ErbB receptor ligands are likely to have great value. The understanding of Spitz neutralization by Argos that we present here provides new avenues to explore in efforts to identify a human homologue of Argos. The structural lessons also provide clear suggestions for which human proteins might be used as scaffolds for generating new protein therapeutics that sequester aberrantly-produced EGF-like growth factors – exploiting a mechanism for inhibiting EGFR signalling that has evolved naturally.

**Methods Summary**

**Protein purification and crystallization:** $Argos_{217}$ was produced in baculovirus-infected *Spodoptera frugiperda* Sf9 cells, using the amino-terminal BiP signal sequence to direct

secretion of the protein into the medium. The protein has a hexahistidine tag at its C-terminus, which was used for purification as described[5]. The EGF domain of Spitz (residues 48-99: $Spitz_{EGF}$) was generated by proteolytic cleavage of a modified form of secreted full-length secreted Spitz produced in transfected *Drosophila* S2 cells. Crystals were grown using the hanging drop method. Crystals grew from a 1:1 $Argos_{217}$:$Spitz_{EGF}$ mixture (250μM complex) at pH 7, using PEG20000 as the precipitant, or from a 250μM solution of $Argos_{217}$ alone using PEG3350 as precipitant (at pH 4.5). Crystals of $Spitz_{EGF}$ alone (500μM) grew in ammonium sulphate, pH 6.5.

**Structure determination:** The $Argos_{217}$:$Spitz_{EGF}$ structure was determined by multiwavelength anomalous dispersion (MAD) using the halide soak method[10]. Crystallographic data were collected at the Advanced Photon Source (Argonne, IL) and the Advanced Light Source (Berkeley, CA), as summarized in Supplementary Table 1. Phasing from 10 ordered bromine ions yielded a readily interpretable electron density map allowing nearly the entire chain of each complex to be traced. Alternating cycles of model building with COOT[28] and refinement with REFMAC[29] led to a complete model of Argos and Spitz with $R_{cryst}$ and $R_{free}$ values of 0.20 and 0.24 respectively to 1.6Å resolution (Supplementary Table 1). The unliganded Argos structure was solved by sequential molecular replacement using PHASER in the CCP4 suite of programs[29], and the $Spitz_{EGF}$ structure was solved by molecular replacement using a loop-truncated version of the human EGF domain structure (1JL9)[30].

**Full Methods** and any associated references are available in the online version of the paper at www.nature.com/nature.

1. Holbro, T. & Hynes, N. E. ErbB receptors: directing key signaling networks throughout life. *Annu. Rev. Pharmacol. Toxicol.* **44,** 195-221 (2004).

2. Shilo, B. Z. Regulating the dynamics of EGF receptor signaling in space and time. *Development* **132,** 4017-4027 (2005).

3. Hynes, N. E. & Lane, H. A. ERBB receptors and cancer: the complexity of targeted inhibitors. *Nature Rev. Cancer* **5,** 341-354 (2005).

4. Freeman, M., Klambt, C., Goodman, C. S., & Rubin, G. M. The argos gene encodes a diffusible factor that regulates cell fate decisions in the Drosophila eye. *Cell* **69,** 963-975 (1992).

5. Klein, D. E., Nappi, V. M., Reeves, G. T., Shvartsman, S. Y., & Lemmon, M. A. Argos inhibits epidermal growth factor receptor signalling by ligand sequestration. *Nature* **430,** 1040-1044 (2004).

6. Kretzschmar, D. *et al.* Giant lens, a gene involved in cell determination and axon guidance in the visual system of *Drosophila melanogaster*. *EMBO J.* **11,** 2531-2539 (1992).

7. Tsetlin, V. Snake venom α-neurotoxins and other 'three-finger' proteins. *Eur. J. Biochem.* **264,** 281-286 (1999).

8. Greenwald, J., Fischer, W. H., Vale, W. W., & Choe, S. Three-finger toxin fold for the extracellular ligand-binding domain of the type II activin receptor serine kinase. *Nat. Struct. Biol.* **6,** 18-22 (1999).

9. Barinka, C. *et al.* Structural basis of interaction between urokinase-type plasminogen activator and its receptor. *J. Mol. Biol.* **363,** 482-495 (2006).

10. Dauter, Z., Dauter, M., & Rajashankar, K. R. Novel approach to phasing proteins: derivatization by short cryo-soaking with halides. *Acta Crystallogr. D Biol. Crystallogr.* **56(Pt 2),** 232-237 (2000).

11. Zhang, C. & Kim, S.-H. The anatomy of protein β-sheet topology. *J. Mol. Biol.* **299,** 1075-1089 (2000).

12. Holm, L. & Sander, C. Protein structure comparison by alignment of distance matrices. *J. Mol. Biol.* **233,** 123-138 (1993).

13. Krissinel, E. & Henrick, K. Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr. D Biol. Crystallogr.* **60,** 2256-2268 (2004).

14. Allendorph, G. P., Vale, W. W., & Choe, S. Structure of the ternary signaling complex of a TGF-beta superfamily member. *Proc. Natl. Acad. Sci. U. S. A.* **103,** 7643-7648 (2006).

15. Alvarado, D., Evans, T. A., Sharma, R., Lemmon, M. A., & Duffy, J. B. Argos mutants define an affinity threshold for spitz inhibition in vivo. *J. Biol. Chem.* **281,** 28993-29001 (2006).

16. Garrett, T. P. J. *et al.* Crystal structure of a truncated epidermal growth factor receptor extracellular domain bound to transforming growth factor alpha. *Cell* **110,** 763-773 (2002).

17. Ogiso, H. *et al.* Crystal structure of the complex of human epidermal growth factor and receptor extracellular domains. *Cell* **110,** 775-787 (2002).

18. Lawrence, M. C. & Colman, P. M. Shape complementarity at protein/protein interfaces. *J. Mol. Biol.* **234,** 946-950 (1993).

19. Huai, Q. *et al.* Structure of human urokinase plasminogen activator in complex with its receptor. *Science* **311,** 656-659 (2006).

20. Rösel, M., Claas, C., Seiter, S., Herlevsen, M., & Zöller, M. Cloning and functional characterization of a new phosphatidyl-inositol anchored molecule of a metastasizing rat pancreatic tumor. *Oncogene* **17,** 1989-2002 (1998).

21. Temerinac, S. *et al.* Cloning of PRV-1, a novel member of the uPAR receptor superfamily, which is overexpressed in polycythemia rubra vera. *Blood* **95,** 2569-2576 (2000).

22. Hansen, L. V., Laerum, O. D., Illemann, M., Nielsen, B. S., & Ploug, M. Altered expression of the urokinase receptor homologue, C4.4A, in invasive areas of human esophageal squamous cell carcinoma. *Int. J. Cancer* **122,** 734-741 (2008).

23. Kenny, P. A. & Bissell, M. J. Targeting TACE-dependent EGFR ligand shedding in breast cancer. *J. Clin. Invest.* **117,** 337-345 (2007).

24. Zhou, B. B. *et al.* Targeting ADAM-mediated ligand cleavage to inhibit HER3 and EGFR pathways in non-small cell lung cancer. *Cancer Cell* **10,** 39-50 (2006).

25. Fujimoto, N. *et al.* High expression of ErbB family members and their ligands in lung adenocarcinomas that are sensitive to inhibition of epidermal growth factor receptor. *Cancer Res.* **65,** 11478-11485 (2005).

26. Hynes, N. E. & Schlange, T. Targeting ADAMS and ERBBs in lung cancer. *Cancer Cell* **10,** 7-11 (2006).

27. Borrell-Pagès, M., Rojo, F., Albanell, J., Baselga, J., & Arribas, J. TACE is required for the activation of the EGFR by TGF-alpha in tumors. *EMBO J.* **22,** 1114-1124 (2003).

28. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60,** 2126-2132 (2004).

29. CCP4 (Collaborative Computational Project Number 4) The CCP4 suite: Programs for protein crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **50,** 760–763 (1994).

30. Lu, H. S. *et al.* Crystal structure of human epidermal growth factor and its dimerization. *J. Biol. Chem.* **276,** 34913-34917 (2001).

**Supplementary Information** is linked to the online version of the paper at www.nature.com/nature.

**Acknowledgements**

We thank members of the Lemmon and Ferguson laboratories, Greg Van Duyne, and Jim Shorter for advice and critical reading of the manuscript. Supported by grants from the NIH (to M.A.L.) and U.S. Army Breast Cancer Research Program (to D.E.K. and M.A.L.).

**Author Contributions**

D.E.K. and M.A.L. conceived and designed the project. D.E.K. was responsible for all construct design and execution of protein biochemistry, crystallization, and data collection. D.E.K. solved and refined the $Argos_{217}$/$Spitz_{EGF}$ complex structure. S.E.S. solved and refined the structures of uncomplexed $Argos_{217}$ and $Spitz_{EGF}$ by molecular replacement using datasets collected by D.E.K. K.N. helped with crystal manipulation and data collection. F.S. performed binding studies with Argos and Spitz variants, as well as analytical ultracentrifugation, directed by D.E.K. D.E.K. and M.A.L. interpreted data and wrote the manuscript.

**Author Information**

Coordinates have been deposited in the Protein Data Bank under codes 3CA7 ($Spitz_{EGF}$), 3C9A ($Argos_{217}$/$Spitz_{EGF}$ complex), and 3CGU ($Argos_{217}$ alone). Reprints and permissions information is available at www.nature.com/reprints. The authors declare that they have no competing financial interests. Correspondence and requests for materials should be addressed to M.A.L. (e-mail: mlemmon@mail.med.upenn.edu).

## Online Methods

**Argos constructs.** To establish the normal signal sequence cleavage site of *Drosophila melanogaster* Argos, the amino terminus of the mature recombinant protein was sequenced. The amino-terminal sequence was TRLPLEVF, indicating that mature Argos is a 419 amino acid secreted protein. The non-conserved amino terminus of Argos has little predicted secondary structure and contains multiple O-linked glycosylation sites. Fusing a BiP signal sequence to R88 of mature Argos produced a well-behaved protein that did not appear to be O-glycosylated.

*D. melanogaster* Argos also contains a proteolytically labile 120 amino acid insertion of low conservation (compared to other drosophilids), and little predicted secondary structure, between the $4^{th}$ and $5^{th}$ cysteines of the protein. Non-drosophilid Argos homologues contain only a 5 amino acid linker in this region. We replaced the 120 amino acid linked linker of *D. melanogaster* with the corresponding 5 amino acids (PDGRT) found in *Apis mellifera* Argos (Supplementary Fig. 1). The resulting protein ($Argos_{217}$) was well expressed, and was resistant to proteolytic degradation. It contains 217 amino acids, corresponding to residues 88-139 of mature *D. melanogaster* Argos linked (via the PDGRT sequence) to residues 260-419. A hexahistidine tag was appended to the carboxyl terminus to aid purification.

**Argos production and purification.** $Argos_{217}$ used for crystallization of the Argos/Spitz complex was produced by secretion from Sf9 (*Spodoptera frugiperda*) cells using the Bac-to-Bac baculovirus expression system (Invitrogen Inc.) as recommended by the manufacturers. Approximately 3 days after infection of cells with recombinant virus, conditioned Sf900II media (Invitrogen-Gibco) was harvested and separated from cellular material by brief centrifugation. The medium was then passed over TALON resin

(ClonTech Inc.) for immobilized metal affinity chromatography (IMAC). The column was washed with 3-6 volumes of 10mM MES (pH 6.3), 150mM NaCl containing 50mM imidazole. $Argos_{217}$ was subsequently eluted with 300mM imidazole in the same buffer. The eluted protein (>90% pure by Coomassie staining) was directly loaded onto a cation exchange column (S2, Bio-Rad Inc.) in the same buffer, and eluted with a gradient of NaCl concentration ($Argos_{217}$ elutes at ~1M NaCl). Immediately prior to crystallization or binding studies, $Argos_{217}$ was gel-filtered into 10mM MES (pH 6.3), 150mM NaCl on a Superose 12 column (GE Healthcare).

Crystals of unliganded $Argos_{217}$ were obtained with protein produced from Drosophila Schneider 2 (S2) cells as described previously[5], and purified exactly as described above. This protein behaves identically in all respects to $Argos_{217}$ produced by Sf9 cells. Biosensor studies (Supplementary Fig. 2) established that histidine-tagged $Argos_{217}$ binds to $Spitz_{EGF}$ with the same affinity as reported for wild type $Argos_{419}$ in our previous studies[5].

**Production of Spitz EGF domain.** The coding region for the *D. melanogaster* Spitz extracellular region (ending at residue 99) was subcloned into the S2 cell expression vector pMT/BiP/V5-HisA (Invitrogen) so that the sequence RHHHHHHSMSGT immediately follows the BiP signal sequence cleavage site. The first serine in this sequence corresponds to $S_{16}$ of mature secreted Spitz. A Factor Xa cleavage site was also engineered between residues 47 and 48 of secreted Spitz ($N_{46}I_{47}T\underline{IEGR}$ / $T_{48}F_{49}P_{50}$), where $T_{48}$ represents the first residue of the EGF domain. Cleavage with Factor Xa allows removal of the highly glycosylated Spitz amino terminus. In addition, deletion of the amino-terminal 15 amino acids avoids lipid modification of the first cysteine[31], and substantially increases protein yield. S2 cells that stably express this modified form of secreted Spitz were selected by

cotransfection with pCo-PURO, and puromycin selection[32], and the secreted protein was purified exactly as described[5]. Following purification, the protein was cleaved with Factor Xa, and the 52 amino acid EGF domain of Spitz (T48-D99: $Spitz_{EGF}$) was isolated by size exclusion chromatography on a Superdex Peptide column (GE Healthcare) and the N-terminal fragment (plus uncleaved protein) was removed using IMAC. $Spitz_{EGF}$ binds $Argos_{419}$ and $Argos_{217}$ with the same affinity as the intact secreted form of Spitz[5].

**Crystallization.** $Argos_{217}$/$Spitz_{EGF}$ complex crystals grew from 0.1M HEPES pH 7.0 and 24% ethylene glycol at 21°C, with the addition of low concentrations (0.1-1%) of PEG20000 to slow crystal growth and thus improve crystal size and quality. Brief manipulation freed single crystal fragments that grew further over 7 days and were subsequently frozen directly in liquid nitrogen. Maximum single crystal dimensions reached 150μm x 100μm x 50μm. Crystals were of space group P1, with unit cell dimensions: a=50.0Å, b=51.3Å, c=70.0Å and $\alpha$=84.2˚, $\beta$=74.8˚, $\gamma$=75.7˚. There are 2 complexes per asymmetric unit, with a Matthews coefficient of 2.6$Å^3$/Da giving a solvent content of 53%.

Crystals of uncomplexed $Spitz_{EGF}$ grew from 15mM ammonium sulphate in 0.1M MES pH 6.5 containing 24% ethylene glycol. Crystals grew as single rods over 2 weeks, and were frozen directly from the drop in liquid nitrogen. Crystals were of space group C2, with unit cell dimensions: a=58.3Å, b=36.2Å, c=25.4Å and $\alpha$=90˚, $\beta$=103.1˚, $\gamma$=90˚. There is 1 molecule per asymmetric unit, with a Matthews coefficient of 2$Å^3$/Da and a solvent content of 39%.

Crystals of unliganded $Argos_{217}$ grew at 18˚C from 10-20% PEG3350, 0.1M NaAcetate pH 4.5, containing 0.2M ammonium sulphate. Crystals were rapidly passed through paraffin oil for freezing. Crystals were of space group C2 with unit cell

dimensions: a=113.6Å, b=64.2Å, c=72.5Å and $\alpha$=90˚, $\beta$=101.6˚, $\gamma$=90˚. There are 2 molecules per asymmetric unit, with a Matthews coefficient of 2.5Å$^3$/Da and solvent content of 52%.

**Structure Determination.** For experimental phasing, efforts to introduce a variety of anomalous scatterers were made. Halide soaks[10] were a focus as they have been successful for several other disulphide-rich glycoproteins with few reactive side-chains. Immediately prior to freezing, 1M NaBr (in 5% PEG20000, 0.1M HEPES pH 7.0, 24% ethylene glycol) was directly added (1:1) to the Argos$_{217}$/Spitz$_{EGF}$ complex crystal drops. A 3-wavelength MAD data set was collected on a single NaBr-soaked crystal at APS 23-IDD (Argonne, IL). Data on a second NaBr-soaked crystal were collected at a 4$^{th}$ wavelength with less attenuation for higher resolution data. Data were processed with HKL2000[33] and the phases were determined with SHELX C/D/E[34,35], utilizing all 4 data sets and anomalous signal from 10 bromide ions. The resulting electron density map was readily interpretable, allowing almost the entire chain of each complex in the asymmetric unit to be traced straightforwardly. Alternating cycles of model building with COOT[28] and refinement with REFMAC[29] led to a complete model of Argos and Spitz. The first 10 residues in both Argos molecules are not seen in the crystal structure, and nor are the C terminal hexahistidine tags. In addition, the first and last two residues in Spitz$_{EGF}$ could not be located in the complex. NCS averaging was used for initial rounds of refinement but released in the final stages of refinement.

The structure of unliganded Argos$_{217}$ was solved by sequential molecular replacement using PHASER[36] in the CCP4 program suite[29]. Domains 1 and 2 of Argos$_{217}$ from the complex were used to find a molecular replacement solution, and a solution was then

identified for domain 3. The structure of $Spitz_{EGF}$ was solved by molecular replacement using a loop-truncated version of the human EGF domain structure (1JL9)[30].

**Calculations and figure preparation.** Calculations of buried surface were carried out using AREAIMOL in the CCP4 suite of programs[29]. Calculations of surface complementarity ($S_c$)[18] used the program SC in CCP4[29]. Quantitative descriptions of protein domain movement were calculated using the DynDom server[37]. Structure validation was carried out with SFCHECK and PROCHECK in CCP4[29]. Figures were prepared using PyMOL[38].


31. Miura, G. I. *et al.* Palmitoylation of the EGFR ligand Spitz by Rasp increases Spitz activity by restricting its diffusion. *Dev. Cell* **10,** 167-176 (2006).

32. Iwaki, T., Figuera, M., Ploplis, V. A., & Castellino, F. J. Rapid selection of Drosophila S2 cells with the puromycin resistance gene. *Biotechniques* **35,** 482-486 (2003).

33. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Meth. Enzymol.* **276,** 307-326 (1997).

34. Pape, T. & Schneider, T. R. HKL2MAP: a graphical user interface for phasing with SHELX programs. *J. Appl. Cryst.* **37,** 843-844 (2004).

35. Schneider, T. R. & Sheldrick, G. M. Substructure solution with SHELXD. *Acta Crystallogr. D Biol. Crystallogr.* **58,** 1772-1779 (2002).

36. McCoy, A. J., Grosse-Kunstleve, R. W., Storoni, L. C., & Read, R. J. Likelihood-enhanced fast translation functions. *Acta Crystallogr. D Biol. Crystallogr.* **61,** 458-464 (2005).

37. Hayward, S. & Lee, R. A. Improvements in the analysis of domain motions in proteins from conformational change: DynDom version 1.50. *J. Mol. Graph. Model.* **21,** 181-183 (2002).

38. DeLano, W. L., *The PyMOL Molecular Graphics System.* (DeLano Scientific, Palo Alto, CA, USA, 2002).

## Legends to Figures

**Figure 1** Structure of the Argos-Spitz complex.

**a**, Representative experimental electron density (contoured at 1$\sigma$) obtained after MAD phasing, showing a region of the Spitz-binding site on domain 2 (d2A: see Fig 3). The initial model is shown placed in the density. **b**, The same region of an 2Fo-Fc map (contoured at 2$\sigma$) calculated using final phases. The final model is shown placed in the density. In purple, a peak corresponding to a bromide ion is seen in the anomalous difference Fourier map (contoured at 4$\sigma$) using Br peak data. **c**, Cartoon of the $Argos_{217}$:$Spitz_{EGF}$ complex. Domains 1, 2 and 3 are coloured blue, yellow and red respectively. Spitz is green. Disulphide bridges are coloured orange. Two orthogonal views are shown. **d**, Same as in **c**, but with Argos in surface representation. **e**, Cartoon representation of the Spitz EGF domain structure, with A-, B-, and C-loops marked.

**Figure 2** Argos has three similar domains that resemble the 3-finger toxin fold of TGF$\beta$ receptors.

**a**, The overall fold of the three constituent domains in Argos is illustrated using domain 3. The four strands ($\beta$1-$\beta$4) form two fingers (fingers 1 and 2) that resemble those on a left hand (as shown). The 'palm' side of the domain faces out of the page. A knuckle-like protrusion projects below the page. At the top of the domain is a disulphide-bonded core from which emanate the two fingers plus the thumb (marked). Cysteines C1 to C6, which make C1-C3, C2-C5 and C4-C6 disulphides are labelled, as are N- and C-termini. **b**, Domains 1, 2, and 3 are

overlaid (as Cα ribbons) in the same orientation used in **a**. Colours are as in Figure 1. Strand β1', unique to domain 1 is labelled, as is the location of the 120 residue insert (ins) removed to generate $Argos_{217}$. Two orthogonal views are shown. **c**, Domain 3 of Argos (red) is overlaid with the 100aa extracellular ligand-binding domain of the type II activin receptor receptor (ActRII)[14] (coloured light grey: from pdb entry 2GOO).

**Figure 3** Spitz-binding interactions

The centre panel shows the $Argos_{217}$:$Spitz_{EGF}$ complex in an orientation similar to that shown in Fig. 1d (right-hand panel) with Argos coloured grey, and Spitz green. Domains 2 and 3 are marked, as are their two fingers (which project to the left). Four individual Spitz-binding subsites are identified: d2A (yellow), d2B (cyan), d3A (red), and d3B (magenta). Surfaces of side-chains involved in each subsite are coloured accordingly. In each of the four corners, details of an individual subsite are shown, with Argos side-chains coloured for the site. Spitz is green in all panels, and the Argos backbone is grey.

**Figure 4** Argos, EGFR, and structural homologues entrap the EGF domain with two binding sites.

**a**, The left-most and right-most panels show EGF domains bound to Argos and the human EGFR extracellular region[17] (sEGFR) respectively. Spitz is green and hEGF cyan. In the central upper panels, $Spitz_{EGF}$ and hEGF are shown (in

identical orientations) bound to Argos domain 2 (grey) and sEGFR domain I (beige). The side-chains of EGF domain-interacting residues are drawn. Site 1 on sEGFR domain I (defined by Ogiso *et al.*[17]) and its counterpart on Argos (which includes site d2A) are marked by blue and red ovals respectively. In the lower central panels, $Spitz_{EGF}$ and hEGF (again in identical orientations) are shown bound to Argos domain 3 and sEGFR domain III. Sites 2 and 3 in the sEGFR/hEGF interface are marked with blue ovals. Argos site d3B mimics sEGFR site 2, but Argos does not mimic sEGFR site 3. Instead, Argos makes a unique set of interactions with $Spitz_{EGF}$ (site d3A). A key aliphatic side-chain critical for hEGF binding to site 3 of EGFR (L47 in hEGF, I98 in Spitz) is disordered and exposed in the $Spitz_{EGF}$/Argos complex. **b**, Domain organization of the uPA receptor[9,19]. The three domains in uPAR are coloured with the order used for Argos in Fig. 1. Like Argos, uPAR uses three copies of this domain type – although in a different arrangement – to form a C-clamp-like structure for enveloping an EGF domain[9,19].

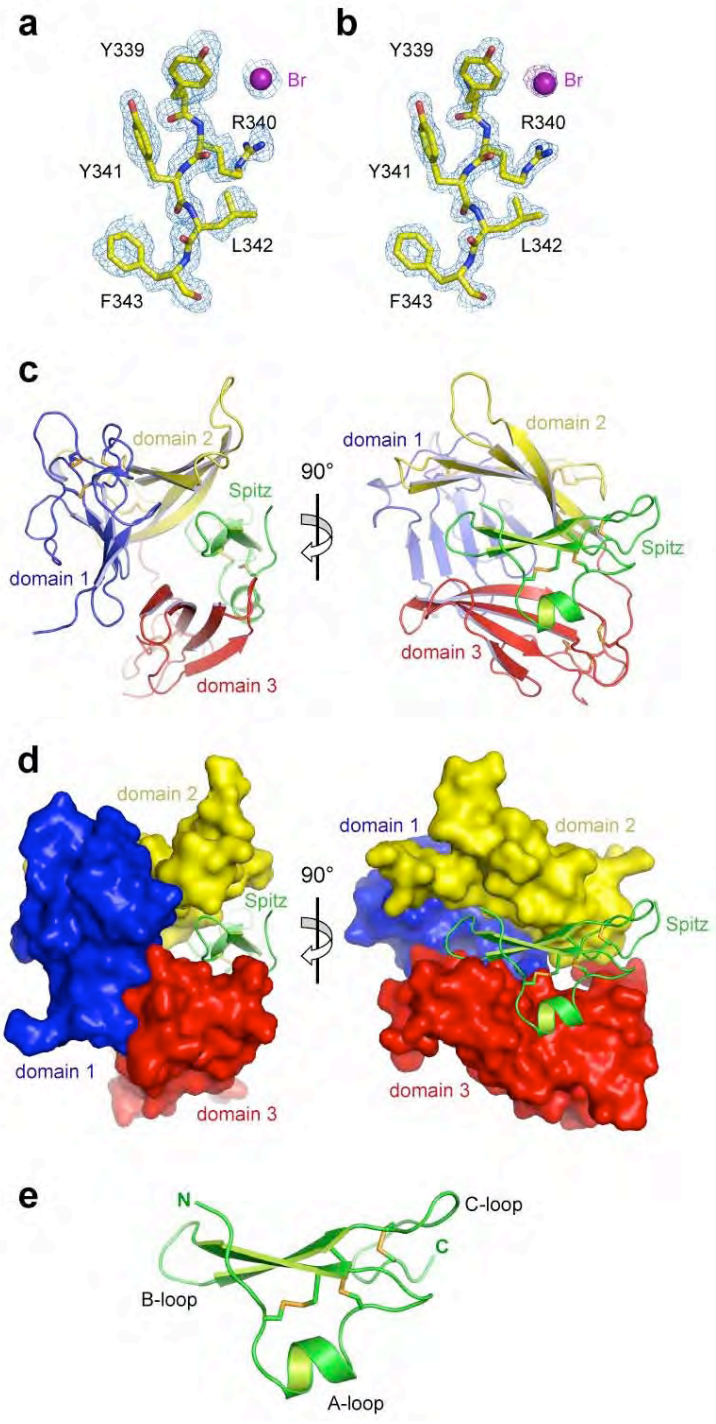
a
Y339
Br
R340
Y341
L342
F343
b
Y339
Br
R340
Y341
L342
F343
c
domain 2
Spitz
domain 1
domain 3
90°
domain 1
domain 2
Spitz
domain 3
d
domain 2
Spitz
domain 1
domain 3
90°
domain 1
domain 2
Spitz
domain 3
e
N
C-loop
C
B-loop
A-loop

Figure 1

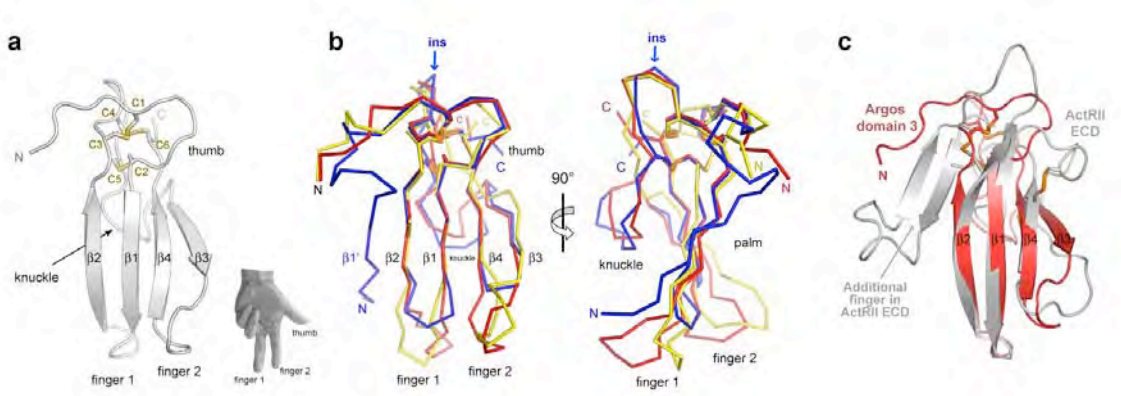


Figure 2

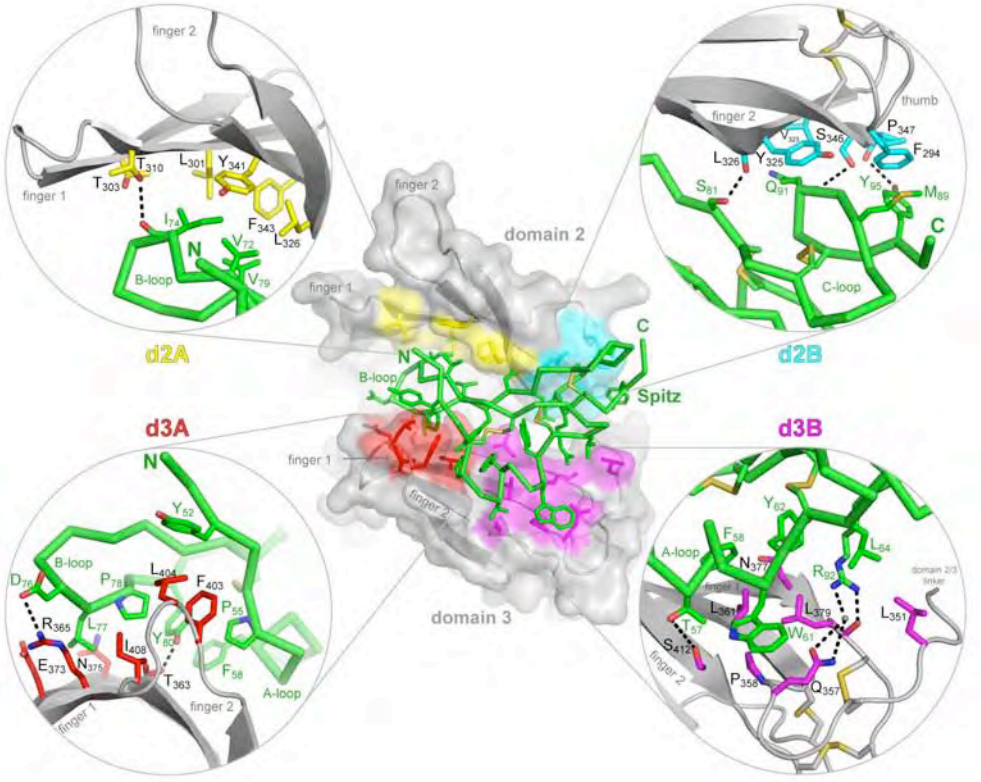


Figure 3

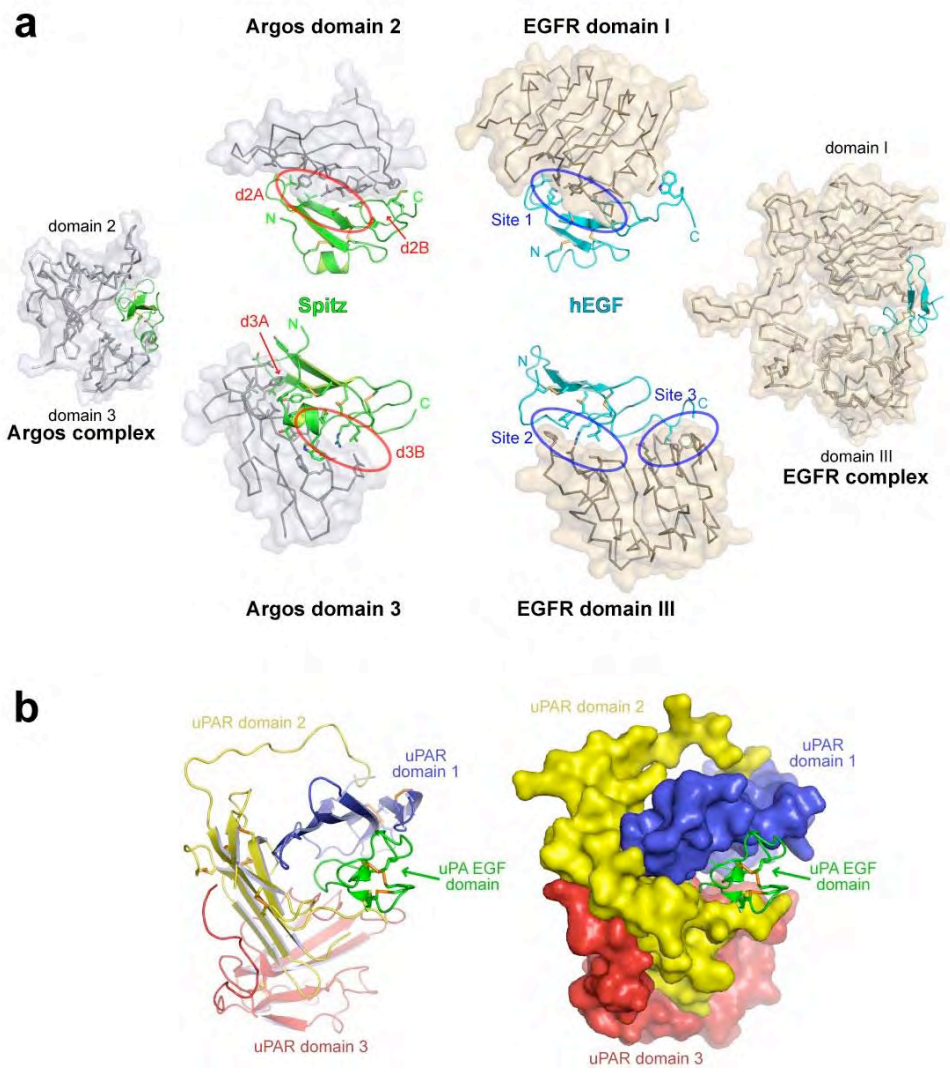
a
Argos domain 2
EGFR domain I
domain 2
d2A
d2B
N
C
Site 1
domain I
Spitz
hEGF
d3A
d3B
Site 3
Site 2
domain 3
Argos complex
domain III
EGFR complex
Argos domain 3
EGFR domain III
b
uPAR domain 2
uPAR domain 1
uPA EGF domain
uPAR domain 3

Figure 4

**Supplementary Results and Discussion** **Nature 01-00681**

# Structural basis for EGFR ligand sequestration by Argos

Daryl E. Klein, Steven E. Stayrook, Fumin Shi, Kartik Narayan, and Mark A. Lemmon

*Department of Biochemistry and Biophysics, University of Pennsylvania School of Medicine, 809C Stellar-Chance Laboratories, 422 Curie Boulevard, Philadelphia, PA 19104-6059, U.S.A.*

## Intermolecular domain/domain interactions in $Argos_{217}$

Although domain 1 of Argos does not directly contact Spitz, it interacts with domains 2 and 3, and may position these two domains (the jaws of the C-clamp) ideally for their simultaneous interaction with a single Spitz molecule. In addition to being covalently linked, domain 1 makes several specific contacts with domain 2. Residues 119-121 in the domain 1 'thumb' make β sheet-like main-chain hydrogen bonds with strand β2 in the first domain 2 finger (Fig. S4a). Interestingly, we previously showed that mutating the conserved valine at position 121 reduces the Spitz-binding affinity of Argos, and impairs its function[1]. Additional interactions that stabilize the domain 1/2 relationship include hydrogen bonds between the side-chain of Q105 (from the domain 1 amino-terminus) and the main chain of the β2-strand in domain 2. The side-chain of D274 (in the tip of the second domain 1 finger) also hydrogen bonds with T324 (in the second finger of domain 2).

Extensive contacts are also seen between domains 1 and 3 of Argos in the complex (Fig. S4b). The tips of the domain 1 fingers touch the edge (strand β2) of the first domain 3 finger, and make several inter-domain hydrogen bonds. Notably, W132 (from finger 1 of domain1) interacts with a backbone carbonyl in the β2 strand of domain 3, and also contacts aliphatic side-chains in this region (including L371). R277 (from domain 1, finger 2) also contacts residues 374-375 in strand β2 of domain 3. Further domain 1/3 contacts include a salt bridge between the side-chains of E134 and R367, plus hydrogen bonds from Y128 and E129 in the tip of the first domain 1 finger to a pair of histidines (H391 and H392) in the domain 3 knuckle. Overall, interactions between these 2 non-contiguous domains buries ~1200 $Å^2$ of relatively polar (51% polar) surface, with a high surface complementarity ($S_c$ = 0.77).

## Conformational change in Argos upon $Spitz_{EGF}$ binding

A comparison of $Argos_{217}$ structures with- and without bound $Spitz_{EGF}$ reveals a large reorientation of domain 3 upon ligand binding (Fig. S5). Domains 1 and 2 remain fixed in the orientation shown in Fig. S4a, but domain 3 is rotated by ~90˚ about an axis in the domain 2/3 linker, and translates 7Å along this axis (Fig. S5). Intriguingly, this conformational change amounts to a 'closure' of the Argos C-clamp in the absence of Spitz. All of the domain 1/3 interactions seen in the complex (Fig. S4b) are lost, and domain 3 'collapses' against domain 2 so that the Spitz binding sites of domains 2 and 3 come into direct contact with one another (right-hand panel in Fig. S5). Comparison of the left-most and right-most panels in Fig. S5 shows that finger 2 of domain 3 replaces the Spitz B-loop on the domain 2 ligand-binding surface.

In the process of binding to unliganded $Argos_{217}$, $Spitz_{EGF}$ must wedge itself between domains 2 and 3, and separate them (proceeding from right to left in Fig. S5). To function as an effective ligand sink, Argos must bind tightly to Spitz – suggesting that the intramolecular

domain 2/3 contacts seen in Fig. S5 do not present a significant impediment to Spitz binding. Indeed, $Argos_{217}$ binds $Spitz_{EGF}$ with $K_D$ = 8nM (Fig. S2), around 20-fold more strongly than Spitz or EGF binding to soluble EGFR extracellular regions[2,3]. The intramolecular domain 2/3 interface in unliganded $Argos_{217}$ has characteristics of a weak interaction, consistent with this expectation. Although an average of 590Å$^2$ (69% apolar) is buried on domains 2 and 3 in this intramolecular interaction (compared with an average of 700Å$^2$ buried by $Spitz_{EGF}$ in each domain), the surface complementarity ($S_c$) is very low (at just 0.48, compared with 0.70 in Argos/Spitz interfaces), suggesting that it is weak and poorly packed.

Unliganded $Argos_{217}$ crystallized as a symmetric dimer, although analytical ultracentrifugation experiments (not shown) reveal that dimerization is rather weak in solution ($K_D \geq 20\mu M$). The crystallographic dimer is stabilized largely by antiparallel association of the β2 strands of two molecules (from finger 1 of domain 3), allowing an 8-stranded β-sheet to continue across the dimer (Fig. S6). Formation of this dimer requires domain 3 to be in the orientation found in crystals of unliganded $Argos_{217}$ (Fig. S5, right-most panel), and domain 3 may be trapped in this position by dimerization in the crystal. Given the feeble interface between domains 2 and 3, it seems likely that the orientation of domain 3 will not be fixed in the $Argos_{217}$ monomer that predominates in solution at physiological concentrations. Spitz binding could then lock a mobile domain 3 into the position seen in the $Argos_{217}$/$Spitz_{EGF}$ complex (which is monomeric in solution[2]).

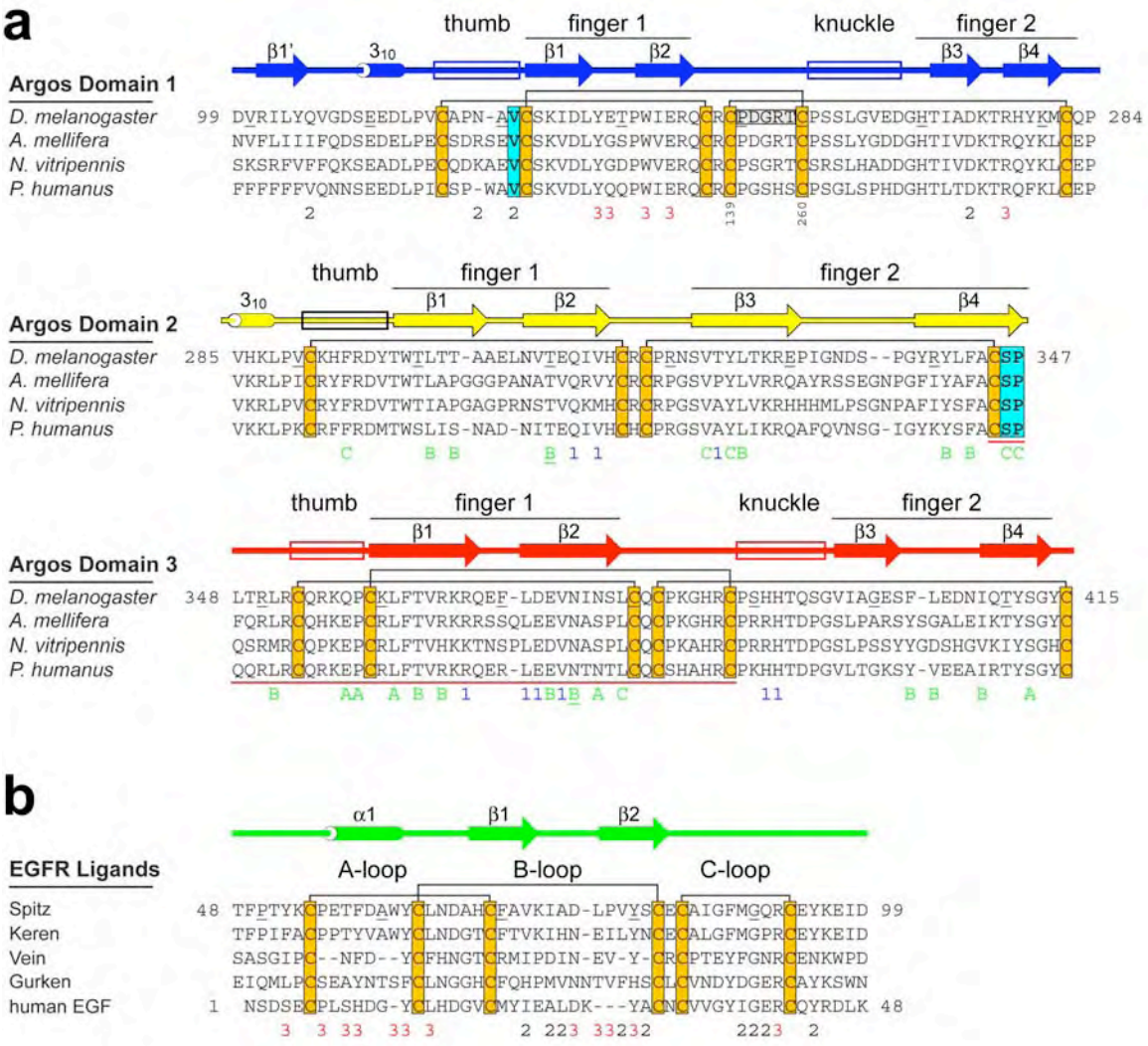


**Figure S1**

**a**, Argos sequences from *Drosophila melanogaster*, *Apis mellifera*, *Nasonia vitripennis* and *Pediculus humanus* are aligned and separated into the three constituent domains. Elements of secondary structure seen in $Argos_{217}$ are indicated above the sequence. In the *D. melanogaster* Argos sequence, every tenth residue is underlined. Positions of the thumb and knuckle loops are boxed (note that domain 2 has no knuckle). Disulphide connectivity is drawn with black lines between cysteines (orange boxes). The PDGRT linker used in $Argos_{217}$ to replace amino-acids 140-259 of intact *D. melanogaster* Argos is boxed grey. Sites at which mutations impaired Argos function in a genetic screen[1] (V121, S346, P347) are boxed with cyan. The initially proposed[4,5] EGF-like domain in Argos (primarily in domain 3) is underlined in red. Residues are marked beneath the alignment according to their inter- and intramolecular interactions. Residues that contact Spitz are labelled (in green) A, B, or C – depending on whether they contact the A-, B- or C-loop of Spitz (see **b**). Argos residues involved in intramolecular domain-domain contacts are numbered according to the domain with which they interact (1, 2 or 3). Two residues involved in both inter and intramolecular contacts (T310 and N375) are underlined. **b**, Alignment of the EGF domains from *D. melanogaster* EGF receptor-activating ligands and human EGF. Disulphides and secondary structure elements are marked as in **a**, and the positions of the A-, B- and C-loops are shown. Beneath the alignment, Spitz residues that contact Argos are labelled 2 (black) or 3 (red), according to the domain in Argos that each contacts.

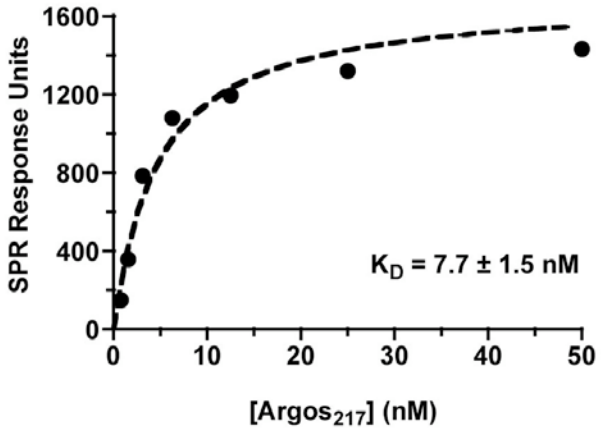


**Figure S2** Binding of $Argos_{217}$ to $Spitz_{EGF}$

Representative surface plasmon resonance (SPR) binding curve for the interaction of soluble $Argos_{217}$ with immobilized $Spitz_{EGF}$. Experiments were performed exactly as described[2]. The mean $K_D$ for $Argos_{217}$ binding to $Spitz_{EGF}$ (7.7±1.5nM) is comparable to values for binding of full-length secreted $Argos_{419}$ to intact secreted Spitz (20±3nM) or $Spitz_{EGF}$ (24±23nM)[2].

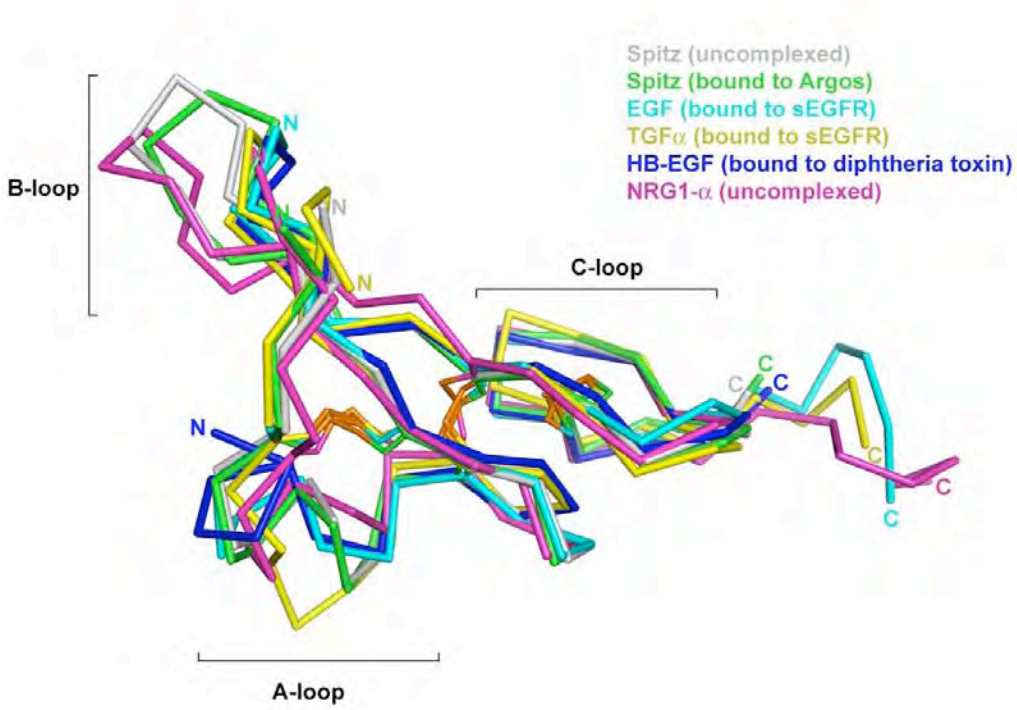


**Figure S3** Overlays of EGF domain structures

Unbound (grey) and Argos-bound (green) Spitz are shown overlaid with crystallographically-derived (sEGFR-bound) structures of EGF[6] (cyan) and TGFα[7] (yellow), diphtheria toxin-bound HB-EGF[8] (blue/purple), and an NMR structure[9] of free NRG1-α (magenta). Note that Spitz most closely resembles NRGα in the length and approximate conformation at the end of the B-loop. The tip of the B-loop is also the only region in which clear differences between free (grey) and bound (green) Spitz can be discerned.

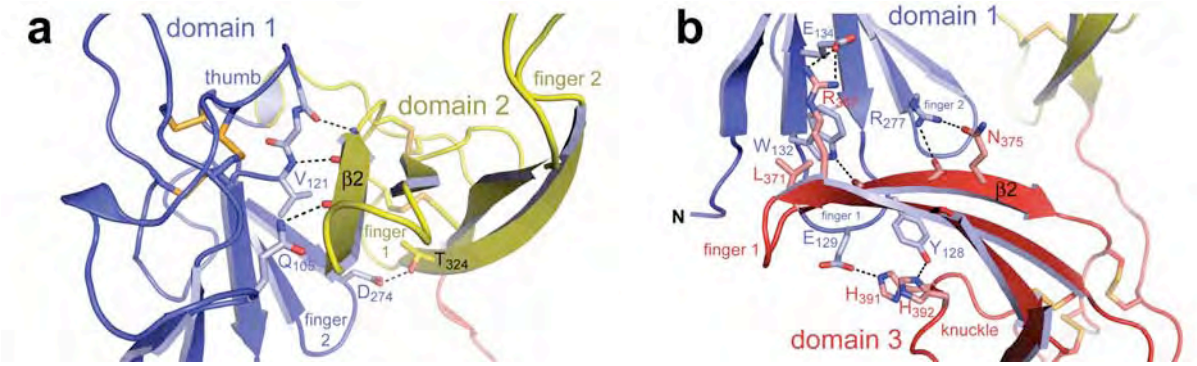


**Figure S4** Interdomain interactions in Argos

**a**, Detail of interactions between Argos domains 1 (blue) and 2 (yellow) in the $Argos_{217}$–$Spitz_{EGF}$ complex. Domain 2 is in approximately the same orientation as in Fig. 1c in main text (left panel). These interactions are not altered by Spitz binding. **b**, Detail of interactions between Argos domain 1 (blue) and 3 (red) in the $Argos_{217}$–$Spitz_{EGF}$ complex. This set of interactions is only seen in the complex. Domains 1 and 3 do not make contact in unliganded Argos.

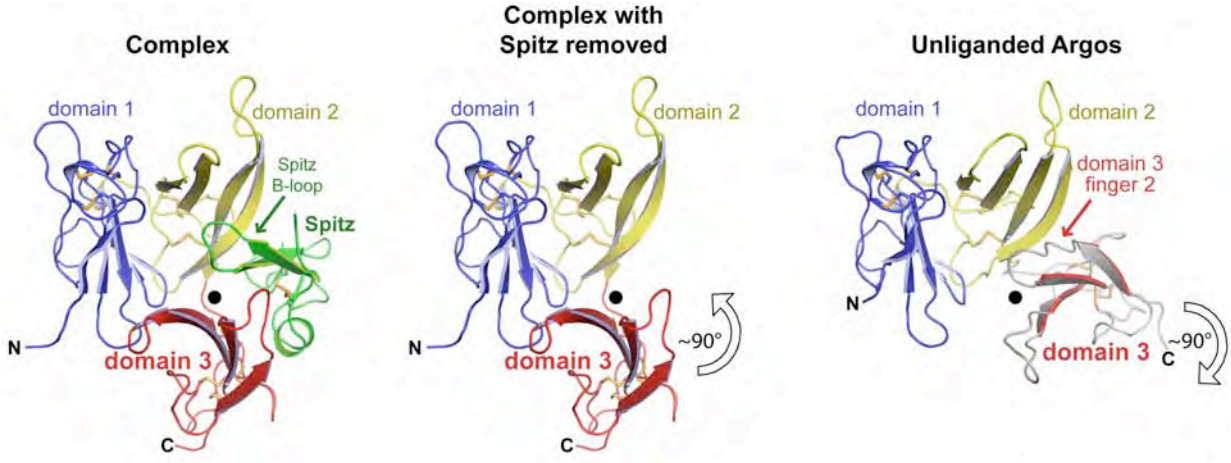


**Figure S5** Conformational change in Argos upon binding to $Spitz_{EGF}$

$Spitz_{EGF}$ binding is associated with reorientation of domain 3. In the absence of bound Spitz, domain 3 of Argos undergoes a rotation of 90˚ about the axis marked with a black circle in the figure. In addition, domain 3 is translated 7Å into the page along this axis. This rotation breaks all domain 1/3 interactions seen in the Argos:Spitz complex (see Fig. S4b). Domain 3 effectively 'collapses' against the Spitz binding site on domain 2, and places its second finger where the Spitz B-loop lies in the complex.

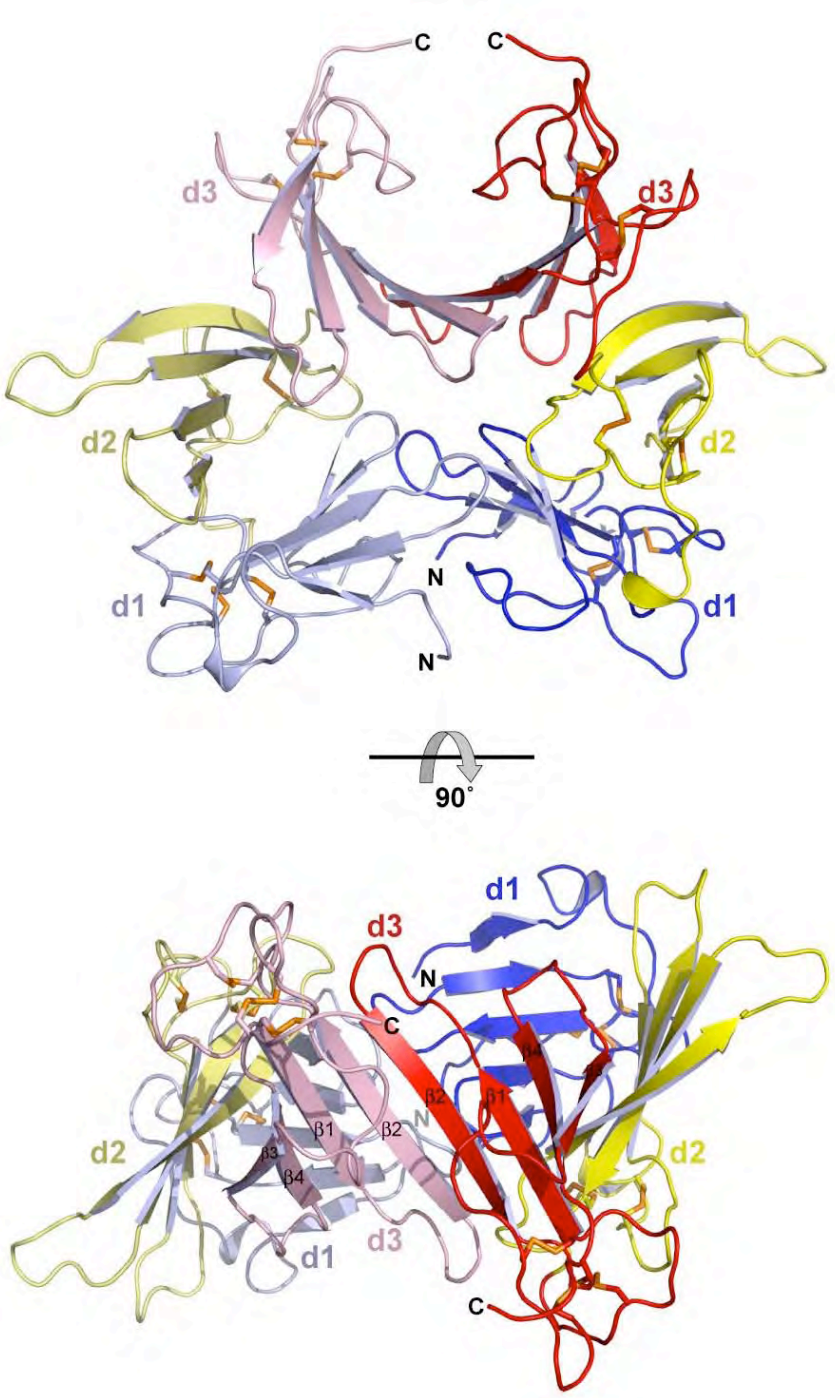


**Figure S6** Dimer of unbound $Argos_{217}$ observed in crystals.

Two orthogonal views are shown for the dimer of $Argos_{217}$ seen in crystals that lack Spitz. Analytical ultracentrifugation studies indicate that this dimer can form in solution, but with a high $K_D$ value of approximately 20μM. Dimerization is mediated primarily by backbone hydrogen bonding that extends the sheet formed by strands β1-β4 of domain 3 across two domains (and two molecules). As described in the text, the domain 3 reorientation that occurs upon Spitz binding prevents formation of this weak dimer.

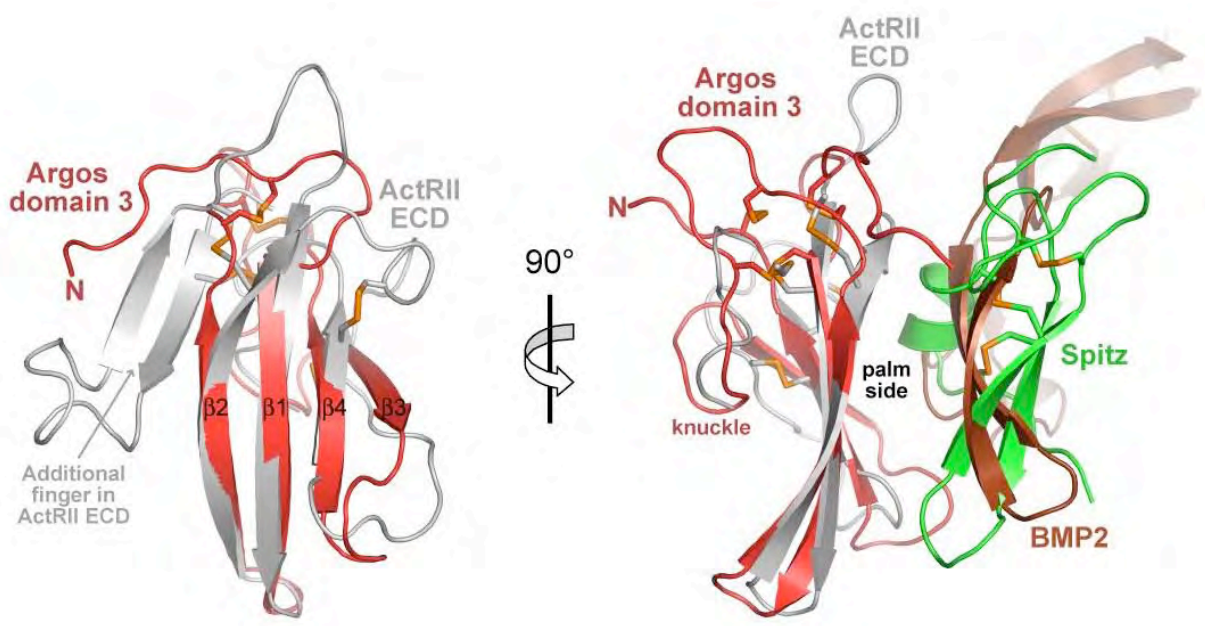


**Figure S7** Similarity of Argos to TGFβ-family receptors

**a**, Domain 3 of Argos (red) is overlaid with the 100aa extracellular ligand-binding domain (ECD) of the type II activin receptor (ActRII)[10] (coloured light grey: from pdb entry 2GOO). The two fingers of each Argos domain overlay well with the longest fingers of the ActRII three-finger toxin fold. In addition, the disulphide-bonded cores are similarly located. The right-hand panel shows an orthogonal view in which Spitz and the receptor-proximal region of BMP2 (bound to ActRII) are shown. Both ligands bind to the palm side of their respective binding domains. The 'third' finger of the ActRII ECD has been removed from this view for clarity. The opposite end of the extended BMP2 molecule binds to a similar site on a type I receptor[10] (not shown).

**References for Supplementary Material**


1. Alvarado, D., Evans, T. A., Sharma, R., Lemmon, M. A., & Duffy, J. B. Argos mutants define an affinity threshold for spitz inhibition in vivo. *J. Biol. Chem.* **281,** 28993-29001 (2006).

2. Klein, D. E., Nappi, V. M., Reeves, G. T., Shvartsman, S. Y., & Lemmon, M. A. Argos inhibits epidermal growth factor receptor signalling by ligand sequestration. *Nature* **430,** 1040-1044 (2004).

3. Ferguson, K. M. *et al.* EGF activates its receptor by removing interactions that autoinhibit ectodomain dimerization. *Mol. Cell* **11,** 507-517 (2003).

4. Freeman, M., Klambt, C., Goodman, C. S., & Rubin, G. M. The argos gene encodes a diffusible factor that regulates cell fate decisions in the Drosophila eye. *Cell* **69,** 963-975 (1992).

5. Kretzschmar, D. *et al.* Giant lens, a gene involved in cell determination and axon guidance in the visual system of *Drosophila melanogaster*. *EMBO J.* **11,** 2531-2539 (1992).

6. Ogiso, H. *et al.* Crystal structure of the complex of human epidermal growth factor and receptor extracellular domains. *Cell* **110,** 775-787 (2002).

7. Garrett, T. P. J. *et al.* Crystal structure of a truncated epidermal growth factor receptor extracellular domain bound to transforming growth factor alpha. *Cell* **110,** 763-773 (2002).

8. Louie, G. V., Yang, W., Bowman, M. E., & Choe, S. Crystal structure of the complex of diphtheria toxin with an extracellular fragment of its receptor. *Mol. Cell* **1,** 67-78 (1997).

9. Jacobsen, N. E. *et al.* High-resolution solution structure of the EGF-like domain of heregulin-alpha. *Biochemistry* **35,** 3402-3417 (1996).

10. Allendorph, G. P., Vale, W. W., & Choe, S. Structure of the ternary signaling complex of a TGF-beta superfamily member. *Proc. Natl. Acad. Sci. U. S. A.* **103,** 7643-7648 (2006).