Dimer Formation through Domain Swapping in the Crystal Structure of the Grb2-SH2−Ac-pYVNV Complex‡

Nikolaus Schiering,*,§ Elena Casale,§ Paolo Caccia,||,⊥ Patrizia Giordano,# and Carlo Battistini#

Departments of Structural Chemistry, Biology, and Chemistry, Pharmacia, Discovery Research Oncology, Viale Pasteur 10, 20014 Nerviano (MI), Italy

Received May 30, 2000; Revised Manuscript Received August 18, 2000

ABSTRACT: Src homology 2 (SH2) domains are key modules in intracellular signal transduction. They link activated cell surface receptors to downstream targets by binding to phosphotyrosine-containing sequence motifs. The crystal structure of a Grb2-SH2 domain−phosphopeptide complex was determined at 2.4 Å resolution. The asymmetric unit contains four polypeptide chains. There is an unexpected domain swap so that individual chains do not adopt a closed SH2 fold. Instead, reorganization of the EF loop leads to an open, nonglobular fold, which associates with an equivalent partner to generate an intertwined dimer. As in previously reported crystal structures of canonical Grb2-SH2 domain−peptide complexes, each of the four hybrid SH2 domains in the two domain-swapped dimers binds the phosphopeptide in a type I β-turn conformation. This report is the first to describe domain swapping for an SH2 domain. While in vivo evidence of dimerization of Grb2 exists, our SH2 dimer is metastable and a physiological role of this new form of dimer formation remains to be demonstrated.

When ligands bind to cell surface receptors, an important event in signal transduction is activation of tyrosine kinases. Src homology 2 (SH2)[1] domains recognize the resulting tyrosine-phosphorylated substrates (*1*). Grb2 is an adapter protein acting upstream of Ras and linking activated receptors to the MAP kinase signaling pathway (*2*). It consists of a central SH2 domain flanked on either side by SH3 domains (*3*). The SH3 domains specifically link Grb2 to the C-terminal tail of the mammalian Ras GDP exchange factors Sos1 and Sos2 (*4*). Direct or indirect binding of Grb2 to activated cell surface receptors recruits Sos to the cell membrane and hence increases the local concentration of Sos for stimulating interaction with membrane-bound Ras (*2*).

Upon activation by HGF/SF, Tyr1349 and Tyr1356 in the carboxy-terminal tail of the hepatocyte growth factor receptor (c-Met) become autophosphorylated and act as a broad specificity docking site for several SH2 domain-containing transducers, including the Grb2−Sos complex. This activation mediates cell growth, motility, and invasion (*5*). c-Met is overexpressed in a variety of human cancers and amplified in certain metastases. c-Met kinase domain mutants identified in hereditary and sporadic forms of papillary renal carcinoma were shown to have elevated enzymatic activity in biochemical assays (*6*). Inhibitors of c-Met downstream signaling, blocking its interaction with the Grb2-SH2 domain, may prove useful as lead compounds for the development of anticancer drugs. Indeed, phosphotyrosine mimetic peptides were found to be active in cellular proliferation and scattering assays (*7*, *8*). A summary of progress achieved in several pharmaceutical laboratories in designing antagonists of SH2 domain interactions is given by Shoelson (*9*). As of today, the most potent small molecule antagonist of the Grb2-SH2 domain is reported by Schoepfer et al. to have a subnanomolar affinity in an ELISA-type assay (*10*).

A large number of SH2 domain structures, both isolated and as part of multidomain proteins, have been reported, determined by either X-ray crystallography or NMR (*11*). The current Protein Data Bank release contains more than 60 structures of more than 16 different SH2 domains, including ligand-bound and ligand-free forms. The fold is conserved among the various SH2 domains with the general scaffold consisting of a central three-stranded antiparallel β-sheet flanked by two α-helices. The central β-sheet is augmented by parallel strands from the N- and C-termini, and a three-stranded β-meander connects strand D with helix B (nomenclature as in ref *12*; see Figure 2A). Peptides bind across the D strand, generally in an extended conformation. A phosphotyrosine binding pocket is formed by residues in the BC loop, αA helix, and B and D strands. Specificity in the SH2−substrate recognition is mediated by interactions

‡ The coordinates have been deposited in the Protein Data Bank as entry 1FYR.

* To whom correspondence should be addressed. E-mail: nikolaus.schiering@eu.pnu.com. Telephone: +39-02-4838 3141. Fax: +39-02-4838 3965.

§ Department of Structural Chemistry.

|| Department of Biology.

⊥ Current address: Ares-Serono, IRCS-Drug Delivery Systems, Via Ribes, 5 10010 Colleretto Giacosa (TO), Italy.

Department of Chemistry.

[1] Abbreviations: Grb2, growth factor receptor-bound protein 2; SH2, Src homology 2; pY, phosphotyrosine (the phosphopeptide numbering is relative to the position of pY which is assigned a 0); SH3, Src homology 3; HGF/SF, hepatocyte growth factor/scatter factor; Sos, Son of sevenless; Stat, signal transducer and activator of transcription; GDP, guanosine diphosphate; GST, glutathione *S*-transferase; Lck, lymphocyte kinase; MAP, mitogen-activated protein; ARP, automated refinement procedure; NCS, noncrystallographic symmetry; NMR, nuclear magnetic resonance; PDB, Protein Data Bank; rms, root-mean-square; RP-HPLC, reversed phase high-performance liquid chromatography; SPR, surface plasmon resonance; ITC, isothermal titration calorimetry.

10.1021/bi0012336 CCC: $19.00
Published on Web 10/14/2000

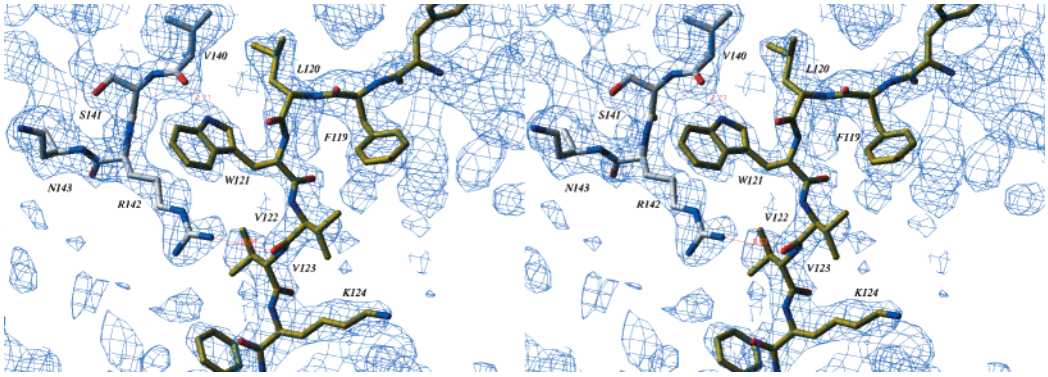

FIGURE 1: Stereoelectron density map of the hinge loop. Shown is a simulated annealing omit map (residues 121−123 and a 3 Å shell around them omitted) at 2.4 Å resolution contoured at 1σ. Chain I is depicted with gray and chain II with yellow carbon atoms; oxygen atoms are in red and nitrogen atoms in blue. This figure was generated using ICM (*47*).

with two to five residues C-terminal to the phosphotyrosine. Selection from degenerate peptide libraries (*13*) and phage display (*14*) have aided in identifying the ideal sequences. The affinities for optimal peptides rank in the submicromolar range. The full-length structure of peptide-free Grb2, including both SH3 domains, was reported in 1995 at 3.1 Å resolution (*15*). Phosphopeptide binding to the Grb2-SH2 domain does not exert a cooperative effect on the SH3 domain and therefore is not expected to influence the Grb2−Sos interaction (*16*). Crystal (*17*−*19*) and NMR (*20*−*22*) structures of the isolated Grb2-SH2 domain were reported, including the unliganded form as well as complexes with peptides and peptidomimetics. Grb2-SH2 is unique in binding the peptide in a β-turn conformation with the specificity-determining residue N+2 pointing toward the SH2 domain surface (*17*).

Here we report the crystal structure of a dimeric form of the Grb2-SH2 domain in complex with the tetrapeptide Ac-pYVNV corresponding to residues 1356−1359 of c-Met.

EXPERIMENTAL PROCEDURES

Peptide Synthesis and Affinity Measurements. The phosphopeptide Ac-pYVNV-OH was synthesized by manual solid-phase synthesis according to standard methods. *N*-Fmoc (9-fluorenylmethoxycarbonyl)-protected amino acids were sequentially coupled starting with the C-terminal Fmoc-Val-OH. After the last coupling step, the tyrosine residue was phosphorylated on the solid phase using phosphoramidite chemistry and successive oxidation. N-Terminal acetylation was obtained by treatment of the peptidyl resin with acetic anhydride. The acetylation was followed by acid treatment leading to side chain deprotection and detachment from the resin. The crude peptide was purified by RP-HPLC (C18 Vydac, Hesperia, CA), achieving a purity of 99.3%.

In an ELISA-type assay, binding of the GST−Grb2-SH2 fusion protein to a biotinylated phosphopeptide containing residues 1345−1360 of c-Met (6-biotinamidohexanoyl-GGGGG-IGEHYVHVNATpYVNVK) was assessed and Ac-pYVNV-OH was tested as an antagonist at increasing concentrations. An IC_{50} of 2.4 μM was obtained.

Expression and Crystallization of Recombinant Grb2 SH2. The c-DNA corresponding to residues 50−161 of human Grb2 was amplified by PCR and cloned into the pGEX-2T expression vector (Pharmacia). The recombinant protein was expressed as a soluble, N-terminal GST fusion protein with a cleavage sequence for thrombin in *Escherichia coli* strain DH5α. The construct includes residues 50−161 of Grb2 and two additional residues at the N-terminus (Gly and Ser; native Grb2 has residues Ile and Pro at positions 48 and 49, respectively). Harvested cells were lysed by sonication, and the fusion protein was bound to glutathione Sepharose (Pharmacia). The SH2 domain was released from the immobilized GST portion through incubation with thrombin and the reaction blocked by addition of hirudin. The SH2 domain was further purified on a S-Sepharose column (Pharmacia). About 9 mg of purified protein were obtained per liter of culture. Both SDS gel electrophoresis after cross-linking and analytical gel filtration (Superdex 75, Pharmacia) showed the presence of the dimer and monomer at a ratio of about 4:1 (data not shown). This protein preparation was concentrated to about 10 mg/mL by ultrafiltration and the phosphopeptide added at a 1.5-fold molar excess. Preliminary crystallization conditions were found in a factorial screen by employing the hanging drop method at 20 °C (*23*). After optimization, crystals grew within a few days by mixing 2 μL of the protein−peptide complex [20 mM Tris-HCl (pH 7.5) and 300 mM NaCl] with the same volume of reservoir solution [11% PEG 3350, 100 mM MES/NaOH (pH 5.7), and 0.5 M NaCl].

Upon further characterization of the Grb2-SH2 preparation, using preparative gel filtration, pure monomer and dimer could be isolated from the mixture. Pure monomer could also be obtained by any of the following: decreasing the pH to 5.0, treatment at 50 °C for 10 min, or including a low percentage of organic solvents such as acetonitrile or 2-propanol in the buffer. Both monomer and dimer are stable at elevated concentrations (up to 10 mg/mL) at pH 7.0 and temperatures of 4, 25, and 37 °C.

Data Collection and Structure Solution. Diffraction images (0.5° oscillation) were collected from two crystals at the X-ray diffraction beamline 5.2R of the Elettra storage ring (Trieste, Italy) at an X-ray wavelength of 1 Å on a Mar-research 180 mm imaging plate detector at distances of 210 and 200 mm, respectively. The crystals were stabilized in mother liquor containing 20% glycerol prior to nitrogen stream freezing. Raw data were reduced with Denzo (*24*). Data for the two crystals were scaled and merged using Scalepack (*24*; Table 1). The structure was determined by molecular replacement (AmoRe from ref *25*) using residues 60−150 of PDB entry 1gri (*15*) as a search model. When all four monomers in the asymmetric unit had been placed, the correlation coefficient was 48.4 and the *R* value 46.7%

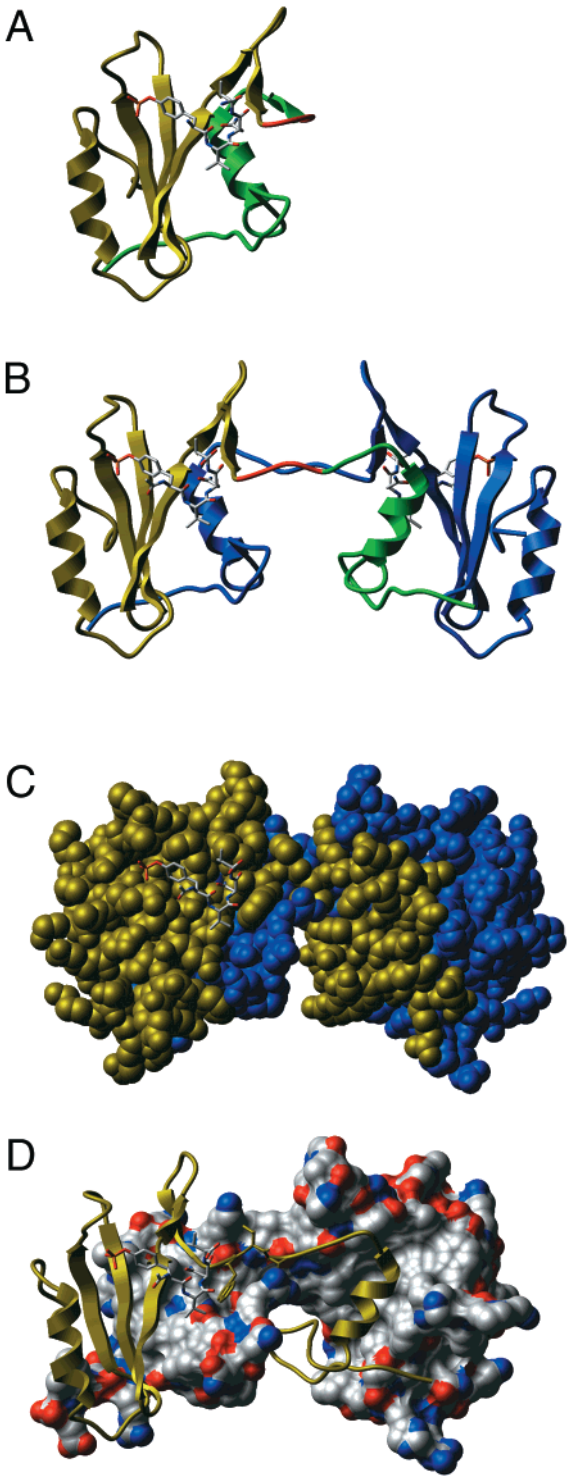

FIGURE 2: Domain swapping generates an intertwined dimer of hybrid SH2 domains. (A) Ribbon diagram of the monomeric, canonical Grb2-SH2 domain structure [PDB entry 1tze (*17*)] with residues 58−120 in yellow, 121−123 in red, and 124−150 in green. Also shown are residues 4−7 of the bound heptapeptide. Throughout this figure, the peptides are shown colored according to atom type (carbon, gray; oxygen, red; nitrogen, blue; and phosphorus, orange). (B) The domain-swapped dimer. Chain I is color coded as described for panel A, and chain II is shown in blue. (C) The domain-swapped dimer in a van der Waals representation with chain I in yellow and chain II in blue. (D) Chain I with residues 121−123, including side chains shown in yellow. Chain II is shown as a surface colored by atom type (carbon, gray; oxygen, red; and nitrogen, blue). The orientation is slightly rotated with respect to the other views. This figure was generated using ICM (*47*).

for data between 10 and 3.5 Å resolution in the correct enantiomorph $P4_32_12$ after rigid-body refinement. The four strongest peaks in a difference Fourier map calculated at this stage of the structure solution corresponded to the predicted positions of the phosphates in the peptides [based on the position in the Lck SH2 domain (*12*)] and confirmed the correctness of the solution.

Refinement. Refinement was performed using X-PLOR (*26*) employing positional refinement and simulated annealing (3000 K). Initially strict NCS constraints were applied, and rebuilding was performed in averaged maps (*27*). At later stages of the refinement, NCS restraints were applied by employing different weights (kilocalories per mole per square angstrom) for protein cores and phosphopeptides (100 for the main chain and 50 for the side chain) and loop segments of residues 88−94, 101−104, 112−117, 121−123, 140−145, and 149 and 150 (5 for the main chain and 0 for the side chain), and nonaveraged maps were used for rebuilding. Water molecules were assigned in difference Fourier maps and with ARP (*28*). Since no connectivity was observed in the EF loop, crystals were redissolved and analyzed using N-terminal sequencing and HPLC/ES mass spectroscopy. The results confirmed the integrity of the protein chain. At this stage of the refinement, it became obvious that a strand exchange with the neighboring SH2 chain was occurring and that two 3D domain-swapped dimers were present in the asymmetric unit. Since the open monomers are formed of two subdomains connected at different angles, two independent segments were defined for the NCS relation. After manual rebuilding that included the correction for domain-swapped dimers, the drop in R_{free} was 1%. In the last X-PLOR refinement round, positional and *B*-value refinement were performed without NCS restraints, resulting in a further drop of both R and R_{free} by 1.6 and 0.4%, respectively (R = 21.9%, R_{free} = 30.5% for data between 6 and 2.4 Å resolution with $F > 2\sigma$). This model was refined further using the maximum likelihood target in CNX [CNS version distributed by Molecular Simulations Inc. (*30*)], including all data and applying a mask bulk solvent correction. Conjugate gradient minimization and individual, restrained *B*-factor refinement were employed without NCS restraints imposed. After a final round of manual rebuilding, including the addition of six residues at the termini and of further water molecules, the *R*-factor is 21.7% and R_{free} 27.0% (Table 1).

The final model contains residues 58−152 (chains I and IV) and 55−152 (chains II and III) as well as the four phosphopeptides. Hence, 70 of 465 residues in the asymmetric unit have not been included in the model and are disordered in our crystal. Procheck analysis (*29*) shows that the structure is of overall good quality; 92.4% of the residues are in the most favored region in the Ramachandran plot and the remainder in the additional allowed region. Solvent accessible surface areas were calculated using a 1.4 Å probe with the MSI SOLVATION module (Delphi and SOLVATION 95.0 User Guide, 1995, Biosym/MSI).

RESULTS AND DISCUSSION

Domain Swapping Leads to the Formation of Dimers of Hybrid SH2 Domains. Domain swapping has been observed in a variety of protein structures and has been proposed as

Table 1: Data Collection and Refinement Statistics[a]

Data Collection				
space group	$P4_32_12$			
unit cell	$a = b = 77.61$ Å, $c = 183.47$ Å			
asymmetric unit	four chains			
V_M (Å^3/D)	2.7 (54% solvent)			
resolution (Å)	20.0−2.4			
no. of total reflections (unique)	140468 (22423)			
completeness (%) (last shell)	98.4 (97.7)			
R_{merge} (last shell)	0.094 (0.252)			
Refinement				
resolution range (Å)	20.0−2.4			
no. of protein atoms	3192			
no. of peptide atoms	168			
no. of water molecules	182			
B-factor model	individual restrained *B*-factor			
R-factor (R_{free})	0.217 (0.270)			
Final Model				
rms deviation from ideal				
bond lengths (Å)	0.007			
bond angles (deg)	1.29			
average *B*-factors (Å^2)				
chain	I	II	III	IV
main chain	36.3	36.5	35.9	37.6
side chain	37.2	37.3	37.0	38.2
B-factor rms deviation (Å^2)				
bonded main chain	0.80			
bonded side chain	0.84			

[a] $R_{merge} = \Sigma_{hkl}\Sigma_i|I_{hkl} - \langle I_{hkl}\rangle|/\Sigma_{hkl}\Sigma_i(I_{hkl})$, where I_{hkl} is the intensity of an individual reflection and $\langle I_{hkl}\rangle$ is the average intensity over symmetry equivalents. *R*-factor $= \Sigma_{hkl}|F_{obs} - F_{calcd}|/\Sigma_{hkl}F_{obs}$, where F_{obs} and F_{calcd} are the observed and calculated structure factor amplitudes, respectively. R_{free} is the equivalent of *R*-factor, calculated for a randomly chosen set of the reflections (5%) that were omitted throughout the refinement process. V_M is the partial specific volume (*46*).

a mechanism for reversible assembly of oligomers from monomers and for the evolution of some oligomeric proteins (*31*). Conformational changes in a hinge loop lead to opening of the monomer, exposing an intrasubunit (C-, closed) interface to the solvent. Two open monomers can dimerize with the swapped domain of one monomer binding to the respective exposed complementary binding surface of the other monomer. Each of the two new intersubunit interfaces corresponds to the C-interface in the monomer. Swapping can occur for entire domains or just a few elements of secondary structure.

In the structure reported here, residues 121−123, forming the EF loop in the canonical Grb2-SH2 domain, fold as an extended chain (Figure 1) and the remaining segment is swapped with a NCS-related partner chain (Figure 2). In this dimeric structure, the individual Grb2-SH2 sequence assumes a noncanonical, open fold (Figures 2 and 3A). Trp121, unique to Grb2 in this position and a key residue influencing the peptide binding mode, is located at the N-terminus of the hinge loop.

About 30% (2500 Å^2) of the total surface area per open monomer is buried upon dimerization, rendering an extensive interaction surface as compared to other oligomeric proteins (*32*). The C-interface is almost entirely hydrophobic with five aromatic (Phe61, -83, -95, and -101 and Trp60) and 12 aliphatic (Val99, -119, -132, and -140, Leu97, -120, -128, -131, and -148, and Ile85, -146, and -151) residues forming the core. Two glutamine residues (positions 106 and 144) as well as Asp113 and Lys124 are involved in electrostatic interactions (Figure 2D).

In addition to the intersubunit contacts corresponding to the C-interface, a novel interaction is a hydrogen bond between the terminal amino group of Arg142 and the carbonyl oxygen of Val122 of the partner chain between open monomers I and II and III and IV (Figure 1). In the crystal structure of the monomeric Grb2-SH2 domain, the side chain of Arg142 is disordered (*17*). It is interesting to note that in the structure of the N-terminal SH2 domain of the phosphatase Syp, Lys91 occupies a position similar to that of Arg142 in Grb2 (*33*).

Four hybrid SH2 domains are present in the asymmetric unit, each composed of two different chains, core residues 60−122 from one chain and swapped domain residues 123−150 from the other chain. The peptide binding sites are accessible on opposite sides of the dimer at a distance of about 25 Å (Cα positions of N+2; distance between phosphates, 44 Å). The hybrid SH2 domains differ only to a limited degree from each other. The rms deviations based on Cα positions range between 0.36 Å (two hybrid SH2 domains composed of chains I and II) and 0.85 Å (two hybrid SH2 domains composed of chains III and IV). Superposition on the canonical Grb2-SH2 structure [PDB 1tze (*17*)] gives rms deviations between 0.42 Å (hybrid composed of the core domain from chain II and the swapped domain from chain I) and 0.67 Å (core from chain IV and the swapped domain from chain III).

Comparison of the Two Dimers in the Asymmetric Unit. The two dimers in the asymmetric unit, I−II and III−IV, differ in the relative orientation of their hybrid SH2 domains. This difference in elbow angles is about 12.5°. A superposition of the four protein chains is shown in Figure 3A. While chain I and chain II are essentially identical, in chain III and chain IV the two subdomains are differently oriented with respect to each other. The rms deviation upon superposition of Cα atoms 60−150 ranges between 0.37 Å (I−II) and 2.15 Å (II−IV).

The Domain-Swapped SH2 Dimer Is Metastable. Analytical gel filtration and cross-linking showed that our purified protein consists of about 80% dimer and 20% monomer. Dimer could be converted to monomer by heat treatment, by addition of organic solvent, or by lowering the buffer pH. Rahuel et al. (*18*) were working with a Grb2 construct including residues 53−161, also expressed as a GST fusion protein. These authors observed dimer formation of the preparation as well and used heat-treated, monomeric protein for crystallization. Consequently, the reported structure has the canonical SH2 fold. Guanidinium hydrochloride extraction of inclusion bodies and refolding of the Grb2 construct 47−168 produced the monomeric protein (*19*). When we incubated either purified monomer or dimer forms at concentrations of ≤10 mg/mL (pH 7.0; 4, 25, and 37 °C), we did not observe interconversion of the two forms. Thus, the energy barrier for dimer−monomer conversion is rather high, and the dimer is stable under mild conditions. These data imply that the dimer observed in our crystal structure is metastable.

Trp121 Can Adopt Two Different Side Chain Rotamer Conformations. Unlike all other wild-type structures of SH2−peptide complexes reported to date, where the peptides bind in extended conformations, pYVNV binds to Grb2-SH2

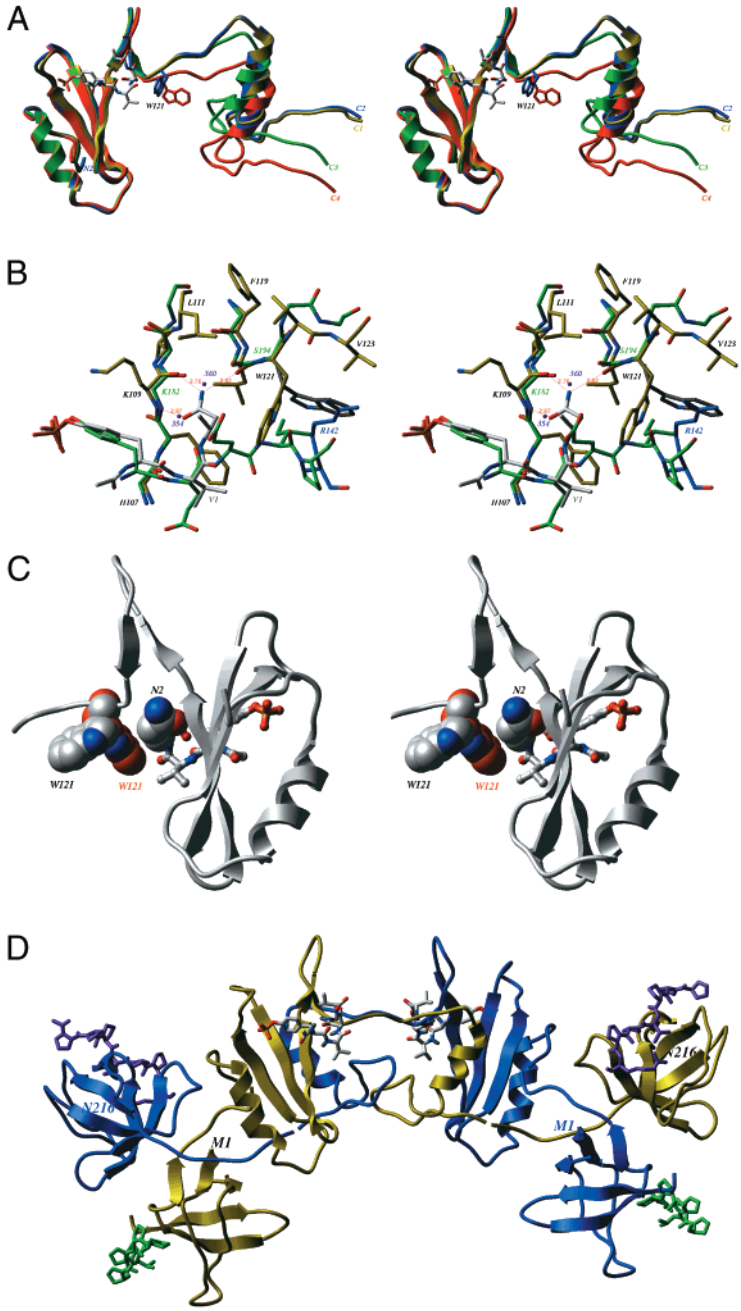

FIGURE 3: (A) Comparison of the four chains in the asymmetric unit. Chain I is shown in yellow, chain II in blue, chain III in green, and chain IV in red (in stereo). The C-termini are indicated for all four chains (C1−C4), and the N-terminus is indicated for chain II (N2). Also shown is the peptide as well as the side chain of Trp121 for chain II and chain IV in blue and red, respectively. The superposition was performed on the core domains (residues 60−120) with root-mean-square deviations ranking between 0.16 Å (II−III) and 0.61 Å (II−IV). (B) Overlay of the SH2−peptide complexes of Grb2 (chain I; carbon atoms in yellow; peptide up to N+2 with carbon atoms in light gray) and Lck [PDB entry 1lcj (*12*); SH2 with the backbone only, carbon in green] as a stereoview. In the case of Lck, two water molecules (354 and 360, shown in purple) anchor the extended peptide via hydrogen bonds to the SH2 backbone. They are replaced by the carboxamide functional group of N+2 in the Grb2−peptide complex. In the superposition shown here, the distance between water 354 and the OD1 atom of the N+2 side chain is 0.6 Å and the distance between water 360 and the ND2 side chain atom is 1.2 Å. Also shown is the alternate conformation of the Trp121 side chain as in chain IV with dark gray carbon atoms as well as Arg142 from chain II with blue carbon atoms. (C) In chains I−III, a stacking interaction between Trp121 and the Cα and Cβ atoms of N+2 of the phosphopeptide contributes to binding affinity by hydrophobic surface burial. Shown are residues 58−125 and the bound peptide of chain IV in gray with the Trp121 and N+2 side chains represented as space-filling diagrams. The Trp121 side chain as found in chain I, representing chains I−III, is shown with orange carbon atoms. The view is rotated by about 180° around a vertical axis in the plane of the paper with respect to panel B. (D) Model of an entire domain-swapped Grb2 dimer. The two Grb2 chains are shown in yellow and blue, respectively, and the bound peptides are colored by atom type (SH2 domain), green (N-terminal SH3 domain), and purple (C-terminal SH3 domain). The N- and C-termini are marked (Met1 and Asn 216). The SH3 domain coordinates were taken from PDB entry 1gri (*15*). The peptides bound to the SH3 domains are based on the X-ray coordinates of the Sem5 C-terminal [PDB entry 1sem (*48*)] and NMR coordinates of the Grb2 N-terminal [PDB entry 1gbr (*49*)] SH3 domains. This figure was generated using ICM (*47*).

as a type I β-turn. Trp121, unique to Grb2, sterically hinders the peptide from assuming an extended conformation. We observe two different side chain conformations for this residue. For chains I−III, χ_1 is between −50° and −56° and

χ_2 is between 83° and 93° [1tze, $\chi_1 = -60°$ and $\chi_2 = 105°$ (*17*)], while for chain IV, the corresponding χ angles are 62° and −67°, respectively. Both conformations correspond to commonly observed Trp rotamers and are shown in panels A−C of Figure 3. While in the former rotamer conformation the indole side chain is oriented toward the bound peptide, with the ring plane roughly normal to the trajectory of an extended peptide chain, in the latter conformation it is swung out by about 100°.

Marengere et al. (*34*) observed a 14-fold reduction in pYVNV affinity of a GST−Grb2-SH2 fusion protein as determined by surface plasmon resonance, when mutating Trp121 to Thr, the corresponding residue in Src-SH2. On the other hand, changing the EF1 residue Thr215 to Trp in Src-SH2 decreased the affinity for pYEEI by 20-fold and increased the affinity for pYVNV by 60-fold to about half of the affinity found with wild-type Grb2-SH2. The affinity for Ac-pYVNV in our binding site IV is likely to be decreased relative to those of the binding sites I−III due to the reduced level of hydrophobic surface burial of the Cα and Cβ moiety of N+2 by the six-membered ring of the Trp121 indole (Figure 3C). Nevertheless, we observe the β-turn binding mode of Ac-pYVNV in binding site IV as in the other three peptide binding sites of the asymmetric unit. The extended peptide binding mode is still sterically excluded by the Trp side chain. The distances between the Ile+3 side chain carbon CD1 and the Trp121 side chain atoms CD1, NE1, and CG in molecule IV range from 1.7 to 2.3 Å in the superposition shown in Figure 3B. ITC measurements using the monomeric and dimeric Grb2-SH2 preparations resulted in equivalent dissociation constants in the low micromolar range for Ac-pYVNV (S. Knapp, personal communication).

Recently, the crystal structure of the Src-SH2 domain EF1 mutant T215W has been reported in complex with a pYVNV-containing dodecapeptide at 2.1 Å resolution (*35*). The β-turn peptide binding mode that is observed is close to identical to that of Grb2-SH2. It is interesting to note that in this structure Trp215, corresponding to Trp121 in Grb2, has a rotamer conformation ($\chi_1 = -62°$ and $\chi_2 = -82°$) different from that of either of the two rotamers reported here. As in our binding site IV, the Trp215 rotamer buries less hydrophobic surface area of N+2 as compared to that observed with the prevailing Trp121 rotamer in Grb2-SH2.

The indole ring of Trp121 packs against the side chain of Arg142′ of the partner chain with either of the two faces of the indole ring being involved in the respective interaction (Figures 1 and 3B). In this way, Arg142 assists in closing the binding groove for an extended peptide chain. In the prevailing Trp121 rotamer, NE1 hydrogen bonds to the carbonyl oxygen of Val140′ (Figure 1), in the one found in binding site IV, to the carbonyl oxygen of Ser139′.

Two Conserved Water Molecules Are Replaced upon Peptide Binding. The β-turn conformation of the peptide ligand allows the side chain of the specificity-determining residue N+2, which in an extended peptide would point to the solvent, to specifically interact with the protein. The Asn side chain is engaged in three hydrogen bonds, to the main chain carbonyl oxygens of Lys109 (βD6) and Leu120 (βE4) as well as to the main chain amide proton of Lys109 (Figure 3B). By interacting with Leu120, N+2 may influence the position of the adjacent residue Trp121 and thereby enforce the conformational limitation of the peptide binding mode. A similar hydrogen bonding network is found in several SH2 domains in complex with extended peptides (*33*, *36*). In these cases, a conserved water molecule links the carbonyl oxygen of peptide residue pY+1 to the amide proton of residue βD6 and via a second bridging water to the carbonyl of βE4. These two water molecules are effectively replaced by the carboxamide group of the Asn side chain in the Grb2-SH2−peptide complex (Figure 3B).

While in conventional SH2−peptide complexes the peptides also interact with the B helix and the BG loop, in the case of Grb2 the peptide contacts are restricted to the core domain (Figure 2C). The peptide binds in a folded conformation with its Asn side chain between the D strand and EF loop, in effect, contacting the break point in the domain swap. The observed domain swapping might reflect not only the dynamics of SH2 domains but also a peculiarity of the pY+2 specificity of Grb2.

Relevance of Dimer Formation through SH2−SH2 Interactions. Cytosolic glutathione *S*-transferases are dimers (*37*), and expression of proteins as GST fusion can favor dimerization. Dimerization of GST-SH2 fusion proteins explains phosphopeptide affinity overestimates as determined by SPR compared to isothermal titration calorimetry (*38*). Panayotou et al. therefore used GST−GST-SH2 heterodimers for SH2−phosphopeptide affinity determination by SPR (*39*). An example of domain swapping being attributed to the expression as a GST fusion protein is dimeric, metastable CD2 (*40*). Deletion of two residues in the hinge loop of CD2 led to a stable form of the intertwined dimer (*41*).

While it is unclear if the dimer formation reported here is physiologically relevant, it is tempting to speculate about a potential role of domain swapping in SH2-mediated dimerization. We therefore generated a model of an entire Grb2 dimer, including the flanking SH3 domains, based on the dimer arrangement in our crystal and the coordinates of full-length Grb2 (*15*; Figure 3D). In the dimer model presented here, the four proline-rich peptides bound to the SH3 domains would be accessible to Sos binding at distances of about 90 Å (distal peptides) and 35 Å (proximal peptides). The two phosphotyrosine phosphate groups bound to the SH2 domain are at a distance of about 44 Å. Harmer and DeFranco (*42*) proposed a model of Ras activation according to which two Grb2 molecules in complex with a single molecule of Shc, an adaptor molecule linking Grb2 to receptors, associate with a single molecule of Sos. They observed that phosphorylation of Shc by the tyrosine kinase Syk generates two Grb2-SH2 binding sites (^{239}pYYN and ^{313}pYVN). In addition, mutagenesis combined with co-immunoprecipitation studies showed that both sites are needed for the efficient formation of Shc−Grb2−Sos complexes in vitro and in vivo.

It is interesting to note that in the structures of dimeric Stat1 (*43*) and Stat3β (*44*) an arrangement of two canonical SH2 domains is observed which corresponds to that of the hybrid SH2 domains in the dimer reported here. Stat proteins dimerize by binding with their SH2 domains to the phosphorylated tail of the other monomer. The SH2 domain of Stat proteins has a rather low degree of sequence similarity with other SH2 domains (about 16% identical with Src), and strands E and F are missing.

Domain swapping has also been observed for the SH3 domain of Eps8 (*45*), expressed as a GST fusion protein. A physiological role as a dimerization motif has been attributed

to the Eps8 SH3 domain since multimerization could be detected in vivo. Analogous in vivo experiments are necessary to test a possible physiological role of SH2 domain swapping for Grb2 signaling.

ACKNOWLEDGMENT

We thank G. A. Petsko (Brandeis University, Waltham, MA) for discussion on domain swapping, S. C. Harrison (Harvard University, Cambridge, MA) for comments on the peptide−water interactions and the manuscript, and W. Kabsch (Max Planck Institute, Heidelberg, Germany) for discussion. We are grateful to G. Lippoli for cloning and L. Piacenza for expression of the Grb2-SH2 domain, L. Sgarella and O. Cletini for protein purification and characterization, D. Volpi for ELISA measurements, S. Knapp for ITC measurements, M. Flocco for critically reading the manuscript, and I. Mathews for discussion. We thank B. Gasperetti (Elettra, Trieste, Italy) for assistance in obtaining the synchrotron beamtime.

REFERENCES

1. Pawson, T. (1995) *Nature 373*, 573−580.
2. McCormick, F. (1993) *Nature 363*, 15−16.
3. Lowenstein, E. J., Daly, R. J., Batzer, A. G., Li, W., Margolis, B., Lammers, R., Ullrich, A., Skolnik, E. Y., Bar-Sagi, D., and Schlessinger, J. (1992) *Cell 70* (3), 431−442.
4. Yang, S.-s., Van Aelst, L., and Bar-Sagi, D. (1995) *J. Biol. Chem. 270*, 18212−18215.
5. Bardelli, A., Pugliese, L., and Comoglio, P. M. (1997) *Biochim. Biophys. Acta 1333* M41−M51.
6. Jeffers, M., Schmidt, L., Nakaigawa, N., Webb, C. P., Weirich, G., Kishida, T., Zbar, B., and Vande Woude, G. F. (1997) *Proc. Natl. Acad. Sci. U.S.A. 94*, 11445−11450.
7. Battistini, C., Giordano, P., De Rosa, S., Corradi, F., Comoglio, P., and Bardelli, A. (1997) International Patent Application WO 97/30079.
8. Gay, B., Suarez, S., Weber, C., Rahuel, J., Fabbro, D., Furet, P., Caravatti, G., and Schoepfer, J. (1999) *J. Biol. Chem. 274*, 23311−23315.
9. Shoelson, S. E. (1997) *Curr. Opin. Chem. Biol. 1*, 227−234.
10. Schoepfer, J., Fretz, H., Gay, B., Furet, P., Garcia-Echevirra, C., End, N., and Caravatti, G. (1999) *Bioorg. Med. Chem. Lett. 9*, 221−226.
11. Kuriyan, J., and Cowburn, D. (1997) *Annu. Rev. Biophys. Biomol. Struct. 26*, 259−288.
12. Eck, M. J., Shoelson, S. E., and Harrison, S. C. (1993) *Nature 362*, 87−91.
13. Songyang, Z., Shoelson, S. E., Chaudhury, M., Gish, G., Pawson, T., Haser, W. G., King, F., Roberst, T., Ratnovsky, S., Lechleider, R. J., Neel, B. G., Birge, R. B., Fajardo, J. E., Chou, M. M., Hanafusa, H., Schaffhausen, B., and Cantley, L. C. (1993) *Cell 72*, 767−778.
14. Gram, H., Schmitz, R., Zuber, J. F., and Baumann, G. (1997) *Eur. J. Biochem. 246*, 633−637.
15. Maignan, S., Guilloteau, J.-P., Fromage, N., Arnoux, B., Becquart, J., and Ducruix, A. (1995) *Science 268*, 291−293.
16. Cussac, D., Frech, M., and Chardin, P. (1994) *EMBO J. 13*, 4011−4021.
17. Rahuel, J., Gay, B., Erdmann, D., Strauss, A., García-Echevírra, C., Furet, P., Caravatti, G., Fretz, H., Schoepfer, J., and Grütter, M. G. (1996) *Nat. Struct. Biol. 3*, 586−589.
18. Rahuel, J., Garcia-Echevirra, C., Furet, P., Strauss, A., Caravatti, G., Fretz, H., Schoepfer, J., and Gay, B. (1998) *J. Mol. Biol. 279*, 1013−1022.
19. Ettmayer, P., France, D., Gounarides, J., Jarosinski, M., Martin, M.-S., Rondeau, J.-M., Sabio, M., Topiol, S., Weidmann, B., Zurini, M., and Bair, K. W. (1999) *J. Med. Chem. 42*, 971−980.
20. Thornton, K. H., Mueller, W. T., McConnell, P., Zhu, G., Saltiel, A. R., and Thanabal, V. (1996) *Biochemistry 35*, 11852−11864.
21. Ogura, K., Tsuchiya, S., Terasawa, H., Yuzawa, S., Hatanaka, H., Mandiyan, V., Schlessinger, J., and Inagaki, F. (1997) *J. Biomol. NMR 10*, 273−278.
22. Senior, M. M., Frederick, A. F., Black, S., Murgolo, N. J., Perkins, L. M., Wilson, O., Snow, M. E., and Wang, Y.-S. (1998) *J. Biomol. NMR 11*, 153−164.
23. Jancarik, J., and Kim, S.-H. (1991) *J. Appl. Crystallogr. 24*, 409−411.
24. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol. 276*, 307−326.
25. Navaza, J. (1994) *Acta Crystallogr. A50*, 157−163.
26. Brünger, A. T. (1992) *X-PLOR version 3.1*, Yale University Press, New Haven, CT.
27. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) *Acta Crystallogr. A47*, 110−119.
28. Collaborative Computational Project, Number 4 (1994) *Acta Crystallogr. D50*, 760−763.
29. Laskovsky, R. A., MacArthur, M. W., Moss, D., and Thornton, J. M. (1993) *J. Appl. Crystallogr. 26*, 283.
30. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr. D54*, 905−921.
31. Bennett, M. J., Schlunegger, M. P., and Eisenberg, D. (1995) *Protein Sci. 4*, 2455−2468.
32. Janin, J., Miller, S., and Chothia, C. (1988) *J. Mol. Biol. 204*, 155−164.
33. Lee, C.-H., Kominos, D., Jacques, S., Margolis, B., Schlessinger, J., Shoelson, S. E., and Kuriyan, J. (1994) *Structure 2*, 423−438.
34. Marengere, L. E. M., Songyang, Z., Gish, G. D., Schaller, M. D., Parsons, J. T., Stern, M. J., Cantley, L. C., and Pawson, T. (1994) *Nature 369*, 502−505.
35. Kimber, M. S., Nachman, J., Cunningham, A. M., Gish, G. D., Pawson, T., and Pai, E. F. (2000) *Mol. Cell 5*, 1043−1049.
36. Nolte, R. T., Eck, M. J., Schlessinger, J., Shoelson, S. E., and Harrison, S. C. (1996) *Nat. Struct. Biol. 3*, 364−373.
37. Ji, X., Zhang, P., Armstrong, R. N., and Gilliland, G. L. (1992) *Biochemistry 31*, 10169−10184.
38. Ladbury, J. E., Lemmon, M. A., Zhou, M., Green, J., Botfield, M. C., and Schlessinger, J. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*, 3199−3202.
39. Panayotou, G., Gish, G., End, P., Truong, O., Gout, I., Dhand, R., Fry, M. J., Hiles, I., Pawson, T., and Waterfield, M. D. (1993) *Mol. Cell. Biol. 13*, 3567−3576.
40. Murray, J. M., Lewis, S. J., Barclay, A. N., and Brady, R. L. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*, 7337−7341.
41. Murray, A. J., Head, J. G., Barker, J. J., and Brady, R. L. (1998) *Nat. Struct. Biol. 5*, 778−782.
42. Harmer, S. L., and DeFranco, A. L. (1997) *Mol. Cell. Biol. 17*, 4087−4095.
43. Chen, X., Vinkemeier, U., Zhao, Y., Jeruzalmi, D., Darnell, J. E., Jr., and Kuriyan, J. (1998) *Cell 93*, 827−839.
44. Becker, S., Groner, B., and Müller, L. (1998) *Nature 394*, 145−151.
45. Kishan, K. V. R., Scita, G., Wong, W. T., Di Fiore, P. P., and Newcomer, M. (1997) *Nat. Struct. Biol. 4*, 739−743.
46. Matthews, B. W. (1968) *J. Mol. Biol. 33*, 491−497.
47. Abagyan, R. A., Totrov, M. M., and Kuznetsov, D. N. (1994) *J. Comput. Chem. 15*, 488−506.
48. Lim, W. A., Richards, F. M., and Fox, R. O. (1994) *Nature 372*, 375−379.
49. Wittekind, M., Mapelli, C., Farmer, B. T., II, Suen, K.-L., Goldfarb, V., Tsao, J., Lavoie, T., Barbacid, M., Meyers, C. A., and Müller, L. (1994) *Biochemistry 33*, 13531−13539.

BI0012336