J Biomol NMR (2011) 51:131–150
DOI 10.1007/s10858-011-9548-7

ARTICLE

Domain cooperativity in multidomain proteins: what can we learn from molecular alignment in anisotropic media?

Tairan Yuwen · Carol Beth Post · Nikolai R. Skrynnikov

Received: 9 May 2011 / Accepted: 7 July 2011

Abstract Many proteins have modular design with multiple globular domains connected via flexible linkers. As a simple model of such system, we study a tandem construct consisting of two identical SH3 domains and a variable-length Gly/Ser linker. When the linker is short, this construct represents a dumbbell-shaped molecule with limited amount of domain–domain mobility. Due to its elongated shape, this molecule efficiently aligns in steric alignment media. As the length of the linker increases, the two domains become effectively uncoupled and begin to behave as independent entities. Consequently, their degree of alignment drops, approaching that found in the (near-spherical) isolated SH3 domains. To model the dependence of alignment parameters on the length of the interdomain linker, we have generated *in silico* a series of conformational ensembles representing SH3 tandems with different linker length. These ensembles were subsequently used as input for alignment prediction software PALES. The predicted alignment tensors were compared with the results of experimental measurements using a series of tandem-SH3 samples in PEG/hexanol alignment media. This comparison broadly confirmed the expected trends. At the same time, it has been found that the isolated SH3 domain aligns much stronger than expected. This finding can be attributed to complex morphology of the PEG/hexanol media and/or to weak site-specific interactions between the protein and the media. In the latter case, there are strong indications that electrostatic interactions may play a role. The fact that PEG/hexanol does not behave as a simple steric media should serve as a caution for studies that use PALES as a quantitative prediction tool (especially for disordered proteins). Further progress in this area depends on our ability to accurately model the anisotropic media and its site-specific interactions with protein molecules. Once this ability is improved, it should be possible to use the alignment parameters as a measure of domain–domain cooperativity, thus identifying the situations where two domains transiently interact with each other or become coupled through a partially structured linker.

Keywords Multidomain proteins · Residual dipolar couplings · Alignment tensor · Generalized degree of order · Domain-domain motion · Flexible linker · Conformational ensemble · Conformational disorder · Tandem SH3 construct · PALES software · Steric alignment · Electrostatic alignment · PEG/hexanol alignment media

Electronic supplementary material The online version of this article (doi:10.1007/s10858-011-9548-7) contains supplementary material, which is available to authorized users.

T. Yuwen · N. R. Skrynnikov (✉)
Department of Chemistry, Purdue University, West Lafayette, IN 47907, USA
e-mail: nikolai@purdue.edu

C. B. Post
Department of Medicinal Chemistry and Molecular Pharmacology, Purdue University, West Lafayette, IN 47907, USA

Introduction

Modular design is a hallmark of eukaryotic proteins. A great number of them contain adaptor domains, such as SH2, SH3, WW, PDZ, FHA, etc., connected through stretches of predominantly flexible linker. These domains present a number of 'generic' binding sites, facilitating assembly of protein complexes that prove to be of immense importance in signal transduction. Binding of adaptor domains to their molecular

targets is typically characterized by relatively low affinity (low μM) and limited specificity. As a consequence, the complexes often have the ability to rapidly dissociate, offering fast dynamic response to external stimuli. Because many domains can bind multiple targets, this mechanism gives rise to sophisticated, non-linear signaling networks, efficiently serving the needs of higher organisms (Cesareni et al. 2005).

As an illustration, let us briefly discuss the adaptor protein c-Crk. In mammalian cells, c-Crk is found in two different isoforms derived from alternative splicing. The smaller species, c-Crk I, are comprised of an SH2 domain followed by a ~15-residue linker and an SH3 domain (N-SH3). The larger form, c-Crk II, additionally contains a ~50-residue linker and an extra SH3 domain (C-SH3) (Kobashigawa et al. 2007).

Through its SH2 domain, Crk binds a variety of phosphotyrosine-containing proteins, most prominently p130Cas. A number of important signaling pathways converge at the p130Cas/Crk assembly, which contains multiple copies of Crk and thus has an ability to bring together a number of different Crk binding partners. The major role of these (interconnected) pathways in cell growth, motility, proliferation, adhesion, and apoptosis has been well documented (Feller 2001). There is also a strong link between Crk-mediated signaling processes and various human cancers, including lung and breast cancers, soft-tissue sarcomas and brain cancers (Linghu et al. 2006; Nishihara et al. 2002; Miller et al. 2003; Takino et al. 2003; Watanabe et al. 2006).

From a structural perspective, one of the key questions that can be asked about Crk and other modular proteins concerns the interactions between domains. On one hand, there is a scenario where two domains, connected through long and flexible linker, behave as essentially independent entities. The domain–domain coupling in this case is minimal. On the other hand, two domains can be significantly coupled through direct interaction and/or a conformationally restricted linker. In the case of Crk adaptor, recent structural study by Inagaki and co-workers showed that c-Crk I conforms more closely to the first scenario, while c-Crk II represents the second type of situation (Kobashigawa et al. 2007). More specifically, in the c-Crk I the two constitutive domains, SH2 and N-SH3, display significant motional freedom, whereas in the c-Crk II the three domains, SH2, N-SH3, and C-SH3, form a compact structure centered around the piece of the inter-SH3 linker. This result highlights the role of the extra domain: the binding surface in Crk C-SH3 has an unusual amino-acid composition, which degrades its ability to bind PxxP ligands; at the same time, C-SH3 occludes the binding site in N-SH3, thus downregulating a number of signaling events (Ogawa et al. 1994; Akakura et al. 2005; Muralidharan et al. 2006; Sarkar et al. 2007; Kobashigawa et al. 2007). Another domain rearrangement, which has the effect of shutting down the Crk-mediated signals, occurs upon phosphorylation of the tyrosine residue in the inter-SH3 linker. The subsequent (intramolecular) binding of the pYxxP motif to the SH2 domain abrogates most of the Crk interactions (Donaldson et al. 2002; Cipres et al. 2007; Peterson and Long 2008; Kobashigawa et al. 2007).

The above example underscores the importance of domain–domain coupling. Note, however, that it is not always possible to investigate domain interactions by straightforward structural methods. Indeed, as already mentioned, these interactions are often dynamic in nature, corresponding to diverse and short-lived conformational species. Two different approaches have been developed to address this situation. The first method relies on backbone ^{15}N relaxation. Indeed, domain cooperativity leads to increases in rotational correlation time τ_R, which can be sensitively detected by relaxation experiments (Zhang et al. 2008; Bae et al. 2009). This effect has been recently explored in-depth by Walsh and co-workers (Walsh et al. 2010) using protein constructs that are similar to those employed in our work. An accurate interpretation, however, can be challenging because spin relaxation is sensitive to many motional modes, including domain–domain motions (Chen and Tjandra 2008; Wong et al. 2009). It can also be difficult to 'calibrate' this method since the existing algorithms for prediction of τ_R have limited accuracy (de la Torre et al. 2000). The second method makes use of residual dipolar couplings. In this approach, one of the domains is loaded with a paramagnetic ion, thus achieving a weak alignment in the external magnetic field. If the two domains are fully coupled, then the second domain shows the same degree of alignment as the first; conversely, if the domains are uncoupled, then the second domain fails to show any alignment (Bertini et al. 2004; Bertini et al. 2007). While elegant and potentially useful, this method faces certain technical difficulties as it depends on suitably engineered metal-binding sites or carefully designed chelate tags (Ikegami et al. 2004; Rodriguez-Castaneda et al. 2006; Keizers et al. 2007).

In this study we turn to more conventional alignment methods involving steric mechanism. We further focus on the important limiting case where a pair of non-interacting domains is connected through a variable-length flexible linker. To model such a system, we manufactured a series of chimera proteins consisting of two copies of chicken α-spectrin SH3 domain (α-spc SH3) (Musacchio et al. 1992) connected through a flexible $(GGSGG)_n$-type linker (Huston et al. 1988). In the case of the short linker, this construct represents a dumbbell-shaped molecule which becomes efficiently oriented in a steric alignment media. Conversely, if the linker is long, then each SH3 domain behaves as an independent entity. Furthermore, since each

individual domain has a nearly spherical shape, it displays only a small degree of alignment. Thus, we expect that the degree of alignment changes in a sigmoidal fashion—from high values (short linker) to low values (long linker). The broad goal of this paper is to quantitate this dependence, both experimentally and by means of simulations. In this manner we seek to establish an important reference case—that of a protein with minimal domain–domain coupling (due to a trivial linker connectivity). Anything beyond this minimal coupling would point toward domain cooperativity. Such cooperativity can arise from domain–domain interactions, an increasingly structured linker region, or a combination of these two factors.

Materials and methods

Plasmid construction

The original pET3d plasmid encoding chicken α-spectrin SH3 domain (full-length construct fSH3) was a gift from Prof. Bernd Reif. Since the N-terminus of α-spc SH3 is highly disordered (Musacchio et al. 1992; Blanco et al. 1997; Chevelkov et al. 2005), we truncated the construct by deleting residues 2–6 (deletion construct dSH3). We have also manufactured three chimera proteins, each of which contained two copies of dSH3 connected via a glycine-serine linker. To link the domains we have used the GSG, GG(GGSGG)GG, and GG$(GGSGG)_5$GG sequences (the latter contains 29 residues). In what follows, these tandem constructs are referred to as dSH3-*sl*-dSH3, dSH3-*ml*-dSH3, and dSH3-*ll*-dSH3, alluding to *s*hort, *m*edium-length, and *l*ong *l*inker, respectively. To produce the cDNA for tandem constructs, two separate PCR reactions were carried out. One PCR product encoded N-terminal dSH3 plus half of the linker region; it was digested with restriction enzymes *NcoI* and *BspEI*. The other PCR product encoded the remaining half of the linker plus C-terminal SH3 domain; it was digested with *BspEI* and *BamHI.* In the next step, the two digested PCR products were cloned into pET3d vector and, at the same time, fused by means of the T4 DNA ligase. Note that the use of *BspEI* sites and T4 DNA ligase does not introduce any unwanted residues in the linker sequence. All plasmids were transformed into BL21(DE3) *E. coli* strain and sequenced to check for accuracy. Primer sequences used to produce the DNA vectors and amino-acid sequences of the SH3 constructs are listed in Table S1.

Protein expression and purification

The bacteria transfected with plasmids were grown in 1 L of M9 minimal media enriched with $^{15}NH_4Cl$ at the temperature 37°C. Protein expression was induced at $OD_{600} \sim 0.6$ by adding 0.5 mM IPTG. After 4 h of incubation the cells were harvested by centrifugation (15 min at 8,000 rpm) and stored at −80°C. The cell pellet was then resuspended in pH 8 lysis buffer and lysed using a French Press. fSH3 was contained in the supernatant; it has been purified using the previously described protocol (Chevelkov et al. 2005). All other constructs formed inclusion bodies during the expression. To purify the precipitated protein material, we first washed the French press lysis pellet four times using 20 mM Tris (pH 8), 1 mM EDTA, 100 mM NaCl, 0.5% Triton X-100 buffer. Each wash was carried out in 30 mL of the buffer (30 min at 18,000 rpm); the last wash was done without the Triton. The pellet was then dissolved in 30 mL of 8 M urea, 20 mM Na-citrate, pH 3.5 denaturing buffer and heated in a water bath at 70°C for 30 min. The solution was subsequently dialyzed over the course of 1 day against 20 mM Na-citrate, pH 3.5 refolding buffer using the dialysis bag with MW cutoff 3 kDa. The refolded protein was purified using the Superdex 75 gel filtration column and then concentrated using the Millipore Amicon devices with 3 kDa MW cutoff. All samples have been analyzed by SDS–PAGE to check the protein molecular weight. Protein concentration determined by UV absorbance at 280 nm was in the range from 0.5 to 1.0 mM. In addition to α-spectrin SH3 domain, we have also prepared two samples of human ubiquitin. The plasmid was supplied by Rachel Klevit laboratory through Addgene repository; the expression and purification procedure was adapted from (Lazar et al. 1997).

NMR Spectroscopy

The experimental measurements involved five constructs of α-spectrin SH3 domain (fSH3, dSH3, dSH3-*sl*-dSH3, dSH3-*ml*-dSH3, and dSH3-*ll*-dSH3) in solution with 20 mM Na-citrate, pH 3.5, 0.02% (w/v) NaN_3. One extra sample of fSH3 additionally included 100 mM NaCl (see Table 1). The oriented samples were prepared by using 5% penta(ethylene glycol) monododecyl ether (C12E5) and 1-hexanol (Rückert and Otting 2000). Two series of samples were prepared with the C12E5-to-hexanol molar ratio $r = 0.85$ and 0.96, respectively. The former condition offers stable alignment around the working temperature 25°C (Rückert and Otting 2000); the latter is the most popular choice in the studies employing PEG-hexanol media (Ulmer et al. 2003; Bewley 2001; Braddock et al. 2002). Special care was taken to maintain the same alignment conditions throughout each series of samples. Toward this goal, we used the stock solutions containing 10% C12E5 and $r = 0.85$ (0.96) fraction of 1-hexanol. The anisotropic samples were prepared by mixing 150 μL aliquot of the stock solution with 150 μL aliquot of the protein solution (to be placed in a Shigemi tube). Each freshly prepared batch of protein was used to make two or three NMR samples, including some replicate

Table 1 Alignment parameters from a series of samples in $r = 0.85$ PEG/hexanol media

	dSH3[a]	dSH3	fSH3	fSH3 (high salt)[b]	dSH3-*sl*-dSH3	dSH3-*sl*-dSH3[c]	dSH3-*ml*-dSH3	dSH3-*ll*-dSH3[a]	dSH3-*ll*-dSH3[c]
A_a (10^{-3})	2.23	3.01	-1.40	-2.27	-1.56	-1.51	0.67	0.65	0.75
R	0.62	0.40	0.37	0.39	0.37	0.36	0.65	0.27	0.29
$\{\alpha,\beta,\gamma\}$[d]	308° 62° 81°	308° 65° 80°	32° 38° 320°	27° 40° 325°	33° 80° 63°	34° 80° 63°	168° 53° 159°	107° 54° 155°	106° 54° 155°
N	29	25	46	27	42	42	39	45	45
Q	0.25	0.30	0.22	0.21	0.22	0.21	0.27	0.26	0.27
$\Delta(^2H)$ (Hz)	24.9	21.8	24.7	25.8	23.3	23.1	16.9	23.5	24.4
GDO (10^{-3})[e]	2.03[f]	2.93	1.19[f]	1.85	1.41[f]	1.37	0.92[f]	0.57[f]	0.63
NSP[g]	1.0	0.98	0.94	0.94	-0.53	-0.53	-0.32	0.26	0.27

The results from replicate samples are shown in shaded columns

[a] The RDC fits are shown in Fig. 6

[b] 100 mM NaCl added to the buffer solution

[c] The results included in Fig. 7a, b (green triangles)

[d] Relative to the crystallographic coordinates 1U06. The rotation as defined in PALES is inverse of the rotation used by other programs (Skrynnikov et al. 2000; Valafar and Prestegard 2004)

[e] Normalized toward $\Delta(^2H) = 20$ Hz

[f] Values used to optimize the electrostatic alignment model

[g] Relative to dSH3

samples. Every sample was allowed to equilibrate in the magnet for 2 h, before the residual quadrupolar coupling from 10% D_2O/90% H_2O solvent, $\Delta(^2H)$, was carefully measured. The $^1H^N$-^{15}N residual dipolar coupling data were collected using the IPAP-HSQC pulse sequence (Yang and Nagayama 1996; Ottiger et al. 1998) on a Varian Inova 600 spectrometer equipped with a triple-resonance probe and z-axis gradient. Data sets for isotropic and partially aligned samples were collected as (192, 576) complex matrices with spectral widths of 1,920 and 9,000 Hz in nitrogen and proton dimensions, respectively. Each experiment took approximately 2 h. After the data were collected, the residual ^2H coupling was re-measured; the $\Delta(^2H)$ drift never exceeded 0.2 Hz. The IPAP spectra were processed using the NMRPipe software package and peak positions were determined by means of the nlinLS fitting routine (Delaglio et al. 1995). In the case of tandem constructs, most of the peaks from the N-terminal domain are neatly overlapped with their counterpart peaks from the C-terminal domain. Hence, the RDC values measured for these residues represent the average of the N- and C-domain couplings. For a small number of sites that give rise to two resolved peaks, we choose to average the experimental RDCs on a pairwise basis (see Experimental Results). The experimental data from the tandem constructs represent, therefore, the average values (listed in Tables S2, S3). In analyzing the results, all datasets have been restricted to the core portion of the SH3 domain, residues 9–60.

Structural models

Each individual α-spectrin SH3 domain was modeled based on the crystallographic structure 1U06 (resolution 1.49 Å) (Chevelkov et al. 2005). Protons and several missing heavy atoms from the side chain of residue D48 were added to the structure and optimized using the facilities of CHARMM (Brooks et al. 2009) and XPLOR-NIH (Schwieters et al. 2003) with the CHARMM22 force field. The resulting domain coordinates were used to build structures of tandem dSH3 and other constructs. The structure 1U06 contains five side chains that are modeled with two alternate conformations. In our main structural model all of them are placed in the conformation found in the chain A of the crystallographic

structure. We have also tested other models, in particular the one where the ionizable E17 side chain is placed in conformation B. The energy of this alternative structure, as evaluated in implicit generalized Born solvent (Dominy and Brooks 1999), was found to be within 6 kcal/mol from the all-A structure. The structure, labeled BAAAA, was therefore deemed suitable for the analyses. In addition, the NMR structure 1AEY (Blanco et al. 1997) was also used.

As an example of the SH3 tandem, let us consider the structural model for dSH3-*ml*-dSH3. The amino acid sequence for this construct is M1(E7…L61)D62-GGGG SGGGG-(E7…L61)D62. The bracketed portions of the sequence correspond to the structured portion of the domains, as represented by the crystallographic coordinates 1U06. The remaining part, comprising the 10-residue linker and single-residue termini, is modeled as a random coil. To generate the coordinates of these segments, we used the program TraDES (routine foldtraj, default input parameters) (Feldman and Hogue 2000). For instance, in order to add N-terminal residue M1 we first generate a random dipeptide with the sequence ME. This peptide is then 'glued' onto the structure of the SH3 domain by overlaying the E residue of the peptide with E7 residue in the SH3 domain (specifically, N, C^{α}, and C' atoms are superimposed in this particular order). Finally, all redundant atoms are deleted. This includes all atoms from the peptide residue E, with the exception of H^{N}. The equivalent procedure is used for the C-terminal residue D62. In this case, we begin with the random dipeptide LD, which is 'glued' to the SH3 domain by overlaying the L residue of the peptide with L61 from the SH3 domain (C', C^{α}, and O atoms are superimposed). Subsequently, all atoms from the placeholder peptide residue L are deleted. The same prescriptions are applied to insert the ten-residue linker segment, DGGGGSGGGG.

The dSH3-*ml*-dSH3 model obtained in this fashion is subjected to a further round of MD refinement (1 ps at 100 K). The goal of the refinement is to relieve some of the strain caused by the 'structure assembly' procedure. The relative position of the domains is preserved during this stage (for this purpose all atoms within the domains have been assigned an artificially large mass, 10 kDa). The refinement protocol is implemented using the torsional angle dynamics option in XPLOR NIH (IVM module (Schwieters and Clore 2001)); the internal coordinates of the domains (all dihedral angles in the residues from 7 to 61, except for the angle φ in residue E7) are fixed and the planarity of peptide planes is enforced. This algorithm is used to produce the ensemble of 10,000 models for dSH3-*ml*-dSH3. Many of these models display severe steric clashes that cannot be resolved by our limited refinement procedure. To weed out these structures, we evaluate the energies of each model and analyze the resulting energy distribution as discussed below.

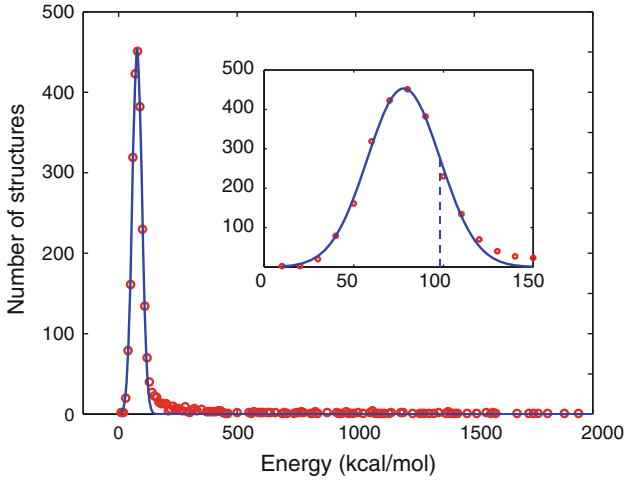

Fig. 1 Part of the energy distribution for 10,000-strong ensemble of dSH3-*ml*-dSH3 models (*red circles*) fitted with the Gaussian function (*blue curve*). The inset displays the expanded portion of the main graph; the *dashed line* corresponds to one standard deviation σ of the Gaussian curve

Figure 1 shows a portion of the energy distribution for 10,000-strong ensemble of dSH3-*ml*-dSH3 models. The energies are evaluated using XPLOR-NIH/CHARMM22 in vacuo (the solvation term is inconsequential, as we seek to identify the structures with severe steric clashes). The low-energy part of the distribution follows the Gaussian curve (see inset in Fig. 1). This is to be expected, since the distribution arises from variability in the random-coil linker region. The Gaussian energy distribution for disordered protein systems can be predicted from general considerations and also has been derived from a number of theoretical models (Bryngelson and Wolynes 1987; Shakhnovich and Gutin 1989; Bryngelson et al. 1995).

In Fig. 1, those structures that comply with the Gaussian distribution constitute a *bona fide* conformational ensemble, whereas the high-energy structures outside the Gaussian distribution (appearing to the right in the plot) suffer from steric conflicts. To address this problem, we analyzed the part of the energy distribution to the right of the dashed line (corresponding to one standard deviation σ of the fitted Gaussian curve) and eliminated a fraction of structures that fall outside the curve. In this manner the original ensemble of 10,000 models was trimmed down to 2,371 models. The retained structures are (1) free of significant steric clashes and (2) have their two domains positioned according to the conformation of the (random-coil) linker, as generated by TraDES. Several randomly selected structures from the resulting reduced ensemble are shown in Fig. 2.

The same procedure as described above has been used to build the models for other chimera proteins, as well as fSH3 and dSH3 (the former carries a 6-residue unstructured

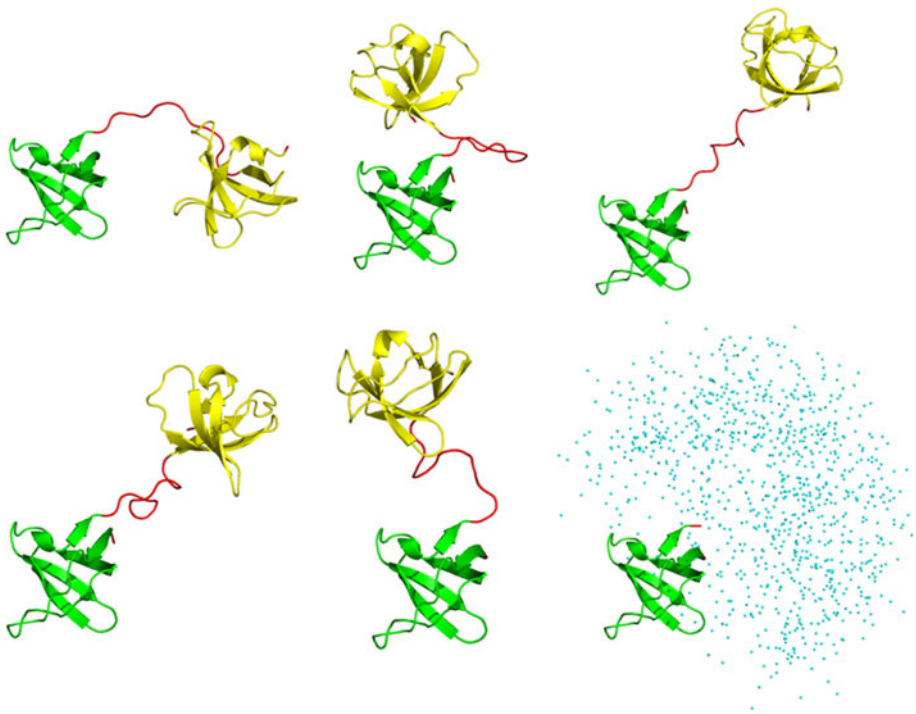

Fig. 2 Five representative structures from the final dSH3-*ml*-dSH3 ensemble. The N- and C-terminal SH3 domains (backbone coordinates same as in the crystallographic structure 1U06, residues 7–61) are plotted in green and gold, respectively; the regions modeled as random coil are colored in *red*. The panel in the lower right corner shows the spatial distribution of the centers of mass of the C-terminal domain in the coordinate frame of the N-terminal domain (as plotted for the entire ensemble)

N-terminus and a single-residue C-terminus, the latter contains two single-residue termini).

Alignment parameters

Protein alignment in liquid–crystalline media is traditionally characterized by the degree of alignment A_a, rhombicity R, and three Euler angles that specify the orientation of the principal axes of the alignment tensor relative to the molecular frame (Tjandra and Bax 1997). These parameters, however, are not necessarily best suited for drawing a comparison between different aligned samples (which becomes especially obvious when Euler angles are compared). For a suitable alternative we turn to the five-dimensional vector space spanned by the irreducible components of the Saupe alignment tensor (Moltke and Grzesiek 1999; Sass et al. 1999; Prestegard et al. 2000). In this linear space the scalar product between the two vectors formed from the Saupe matrix elements can be defined as:

$$\left\langle \mathbf{S}^{sample1} \middle| \mathbf{S}^{sample2} \right\rangle = \sum_{\substack{i=x,y,z \\ j=x,y,z}} S_{ij}^{sample1} S_{ij}^{sample2} \qquad (1)$$

where S_{ij} are the elements of the 3×3 Saupe matrices associated with the rigid molecule or molecular fragment (e.g. SH3 domain within the tandem construct). *sample1* and *sample2* designations refer to samples involving different alignment media or different protein constructs, that share the same fragment (SH3 domain). Note that both $\mathbf{S}^{sample1}$ and $\mathbf{S}^{sample2}$ should be expressed in the same coordinate frame. In our case this frame is linked to the uniquely defined coordinates of the SH3 domain (pdb frame of the structure 1U06).

Using the definition in (1), one can introduce the Generalized Degree of Order (GDO) (Tolman et al. 2001) that characterizes the magnitude of alignment for a given sample:

$$\mathrm{GDO} = \sqrt{(2/3)\left\langle \mathbf{S}^{sample1} \middle| \mathbf{S}^{sample1} \right\rangle} = |A_a| \sqrt{1 + (3/4)R^2}. \qquad (2)$$

Furthermore, if the scalar product (1) is normalized, then it can be viewed as a generalized measure for relative orientation of the two alignment tensors:

$$\mathrm{NSP} = \frac{\left\langle \mathbf{S}^{sample1} \middle| \mathbf{S}^{sample2} \right\rangle}{\sqrt{\left\langle \mathbf{S}^{sample1} \middle| \mathbf{S}^{sample1} \right\rangle \left\langle \mathbf{S}^{sample2} \middle| \mathbf{S}^{sample2} \right\rangle}}. \qquad (3)$$

The values of the normalized scalar product (NSP) close to 1.0 indicate that the two alignment tensors differ only by a scaling factor, whereas the values around 0.0 suggest that the alignment frames are 'orthogonal'. The situation where NSP $= -1$ can be described as 'antiparallel' alignment—the rhombicities and alignment frames are in this case identical, while the degrees of alignment A_a have opposite signs. This situation is observed in stretched versus compressed gels (Sass et al. 2000; Mohana-Borges et al. 2004), conventional versus 'flipped' bicelles (Opella and De Angelis 2007; Koenig et al. 1999), and other similar systems (Denisov et al. 2010). Although NSP defined by

(3) depends also on rhombicity R, it is mainly a measure of alignment axes orientation.

The definition of the Saupe tensor used here (Sass et al. 1999) differs by a factor of 2 from the one used elsewhere (Bax et al. 2001). The definitions of the GDO and NSP, however, are unambiguous. The numeric results for Saupe matrices, as required for evaluating Eqs. (2) and (3), are reported by the program PALES (Zweckstetter and Bax 2000; Zweckstetter 2008), both in the case of the experimentally determined (fitted) and predicted alignment tensors. Finally, the quality factor is defined as (Cornilescu et al. 1998; Bax 2003):

$$Q = \frac{\left(\sum_{i=1}^{N} (D_i^{exptl} - D_i^{calc})^2/N\right)^{1/2}}{\left(\sum_{i=1}^{N} (D_i^{exptl})^2/N\right)^{1/2}}. \quad (4)$$

PALES predictions

The alignment mechanism in PEG/hexanol media is commonly assumed to be steric (Rückert and Otting 2000; Zhang and Zuiderweg 2004; Fischer and Geyer 2005; Maltsev et al. 2008; Berlin et al. 2009). To model this media we used the PALES procedure originally developed for DMPC/DHPC bicelles (Bax and Tjandra 1997). The thickness of the planar bilayer was set to 28.6 Å (Freyssingeas et al. 1996). It is worth mentioning that bicelle calculations in PALES have been programmed for DMPC/DHPC, so that the code implicitly accounts for the presence of 5 mM free DHPC in the solvent (Ottiger and Bax 1998; Simon et al. 2005). To correct for this small contribution and, in addition, account for 0.3 wt% of free hexanol in solution (Freyssingeas et al. 1996) one should use a slightly altered value of the liquid crystal concentration. For instance, in the PALES calculations aimed at 5% PEG, $r = 0.85$ media the effective liquid crystal concentration should be set to 65 mg/ml. The order parameter of the PEG/hexanol bicelle was assumed to be the same (0.8) as for DMPC/DHPC.

We have also attempted to predict the alignment parameters under the assumption that an electrostatic mechanism plays a significant role. The corresponding PALES calculations involve several additional variables: charge density on the surface of the bilayer σ (treated as adjustable parameter), sample pH (actual value 3.5; also treated as adjustable parameter), and ionic strength of the solution (17 mM, corresponding to 20 mM citrate buffer at pH 3.5). To determine the charges on α-spc SH3, we have used the program PROPKA, which calculates the pK_a values of ionizable groups (Li et al. 2005; Bas et al. 2008). Since there is significant structural variability in tandem constructs and fSH3, we have run PROPKA on each individual conformer from the structural ensemble. The results (i.e. the coordinates of a given conformer plus the charges calculated for this particular conformer) were then used as input for PALES calculations. In principle, more sophisticated methods for calculating charges can be used (Gordon et al. 2005; Khandogin and Brooks 2006) and ultimately the pK_a values can be determined experimentally (Tollinger et al. 2002; Andre et al. 2007). However, such effort seems unwarranted given that the models used for prediction of electrostatic alignment remain relatively crude (Zweckstetter et al. 2004; Zweckstetter 2006; Wu et al. 2006). Of note, we have found a programming error in the PALES module responsible for computations of electrostatic alignment. Specifically, natural logarithm is mistakenly used in the Henderson-Hasselbalch equation instead of the base 10 logarithm. As a result, given the set of pK_a (either default or user-supplied values) the program fails to correctly calculate point charges at a given pH. The problem can be circumvented by using a table of user-supplied point charges as the PALES input. In the latest revision of PALES this issue has been corrected (M. Zweckstetter, personal communication, 2011).

Dipolar couplings predicted for the individual conformers are averaged, resulting in a single set of simulated RDCs. When drawing comparison with the experimental data from the tandem constructs, we further average the computed couplings from the N- and C-terminal domains (in accordance with our experimental protocol, see above). Note that such averaging procedure is fully justified so long as both domains have the same (invariant) internal structure. Indeed, RDCs by definition represent the average between many molecular orientations. The resulting dataset is subsequently restricted to the core portion of the α-spc SH3 domain, residues 9–60, and fitted with the (protonated) structure 1U06. In the situation when residues 9–60 are represented by one and the same set of coordinates throughout the course of the calculations, the residual χ^2 from such fitting procedure is strictly zero. The alignment parameters extracted in this fashion are then used to calculate the theoretical values of the GDO and NSP, Eqs. (2) and (3). Note that the averaging of simulated RDCs with respect to the N- and C-terminal domains affects the final GDO and NSP values.

Theory: general outline

Let us consider a generalized two-domain protein, where a pair of globular, near-spherical domains is connected via a flexible structureless linker. In the case when the linker is short, i.e. consists of no more than several residues, the protein has a dumbbell shape. Furthermore, a short linker allows for little motional freedom so that domain–domain motion is highly restricted. To a good approximation, the

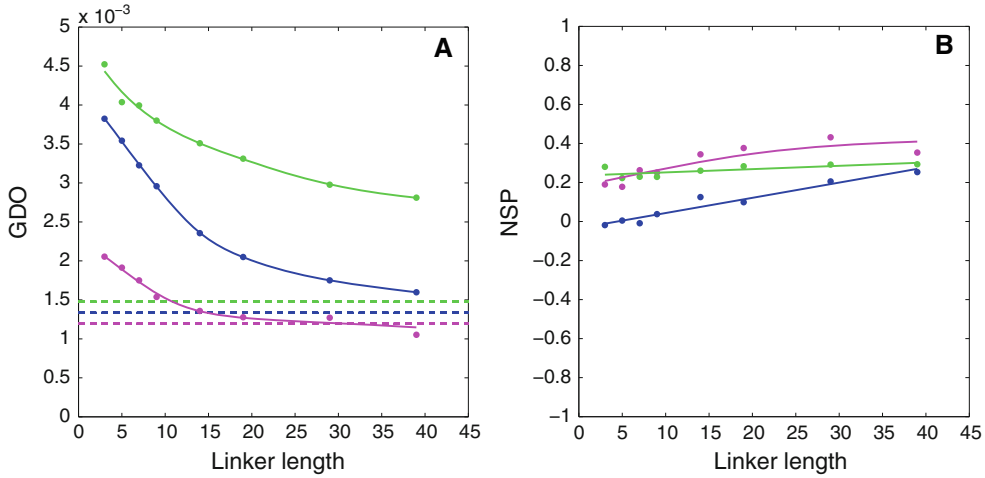

Fig. 3 PALES simulations of steric alignment in a series of computer-generated tandem proteins. To generate structural ensembles, the two identical domains (represented by crystallographic or NMR coordinates and presumed to be fully rigid) were directly connected via glycine/serine linkers. The simulated couplings from all non-proline residues in the N-terminal domain were used to derive the alignment parameters. Other details of the structure generation and RDC simulations are as described in the Materials and Methods. Shown are the results for the tandem constructs of the disintegrin domain (structure 1MPZ (Moreno-Murciano et al. 2003), 41 residue, radius of gyration 8.8 Å, *magenta symbols*), SH2 domain (extracted from structure 1A81 (Futterer et al. 1998), 106 residues, 13.0 Å, *blue*), and nucleotide-binding domain (3GWI (Hakansson 2009), 170 residues, 15.3 Å, *green*). **a** Simulated GDO values for the domains which belong to the tandems (*circles*) and the corresponding isolated domains (*dashed lines*). **b** Simulated NSP values characterizing the mutual projection of the alignment tensors from (1) the domain which belongs to the tandem and (2) the corresponding isolated domain. The spline curves are added to the graphs for visual guidance. To verify the convergence of the simulations, we have regenerated this plot using a different set of randomly generated structural ensembles. The results are shown in Fig. S1

protein can be described as a rigid body with highly elongated shape. Being placed in a steric alignment media, this protein should efficiently align and display a high GDO value, (2). The long axis of the alignment tensor should roughly coincide with the line connecting the centers of mass of the two domains. Generally speaking, the alignment PAS in such a two-domain construct has no relation to the PAS of the individual (near-spherical) domain. Therefore the NSP between the two respective tensors (3), should deviate from 1.0 (unless by coincidence).

Let us now consider the opposite situation, where the flexible linker is very long. In this case the two domains move almost freely and no longer sense each other's presence. Their GDOs, therefore, should approach that of an individual domain. Given that the shape of each domain is nearly spherical, the GDO should drop to a low value. Likewise, the PAS orientation should generally approach that found in the individual domain. Consequently, the scalar product between the two alignment tensors—the one from the domain which belongs to the tandem and the one from the isolated domain—should converge toward 1.0.

Note that neither GDO, nor NSP are expected to show a quantitative convergence toward the single-domain limit. Indeed, the behavior of a tethered domain is similar, but not identical, to that of the isolated domain. In particular, if the shape of the domain is very close to spherical then even a small alteration of the shape (such as caused by attachment of the linker) can lead to a significant change in magnitude and especially orientation of the alignment tensor. Therefore, the convergence toward the single-domain limit should be rather viewed as a broad trend.

To illustrate this expected behavior, we have undertaken a series of PALES simulations for a number of protein domains, Fig. 3. For these simulations we have chosen the domains of different size with close-to-spherical shape; if desired, the results can be extended to domains with elongated or flattened shapes. The coordinates of the domains have been used to generate tandem structures, which were subsequently processed by PALES. As expected, increasing the linker length from three residues to five, then to seven, etc. leads to pronounced decreases in the GDO values. Eventually, each curve approaches a plateau which is close to, but distinct from the GDO of the isolated domain (dashed horizontal lines in Fig. 3a). Similarly, as the linker length is increased, the NSP value becomes closer to 1.0 (Fig. 3b). The trend, however, is weak, which means that the orientation of the alignment frame in the tethered domain remains quite different from the isolated domain.

It is interesting to discuss how the results scale with the size of the system. Generally, one would expect that for the smallest unit, disintegrin, the two domains become 'uncoupled' and the GDO (NSP) plateau is reached already with a moderately long linker. Conversely, for the biggest unit, nucleotide-binding domain (NBD), this does not happen until the linker becomes much longer. A rough theoretical estimate can be readily obtained for the scaling

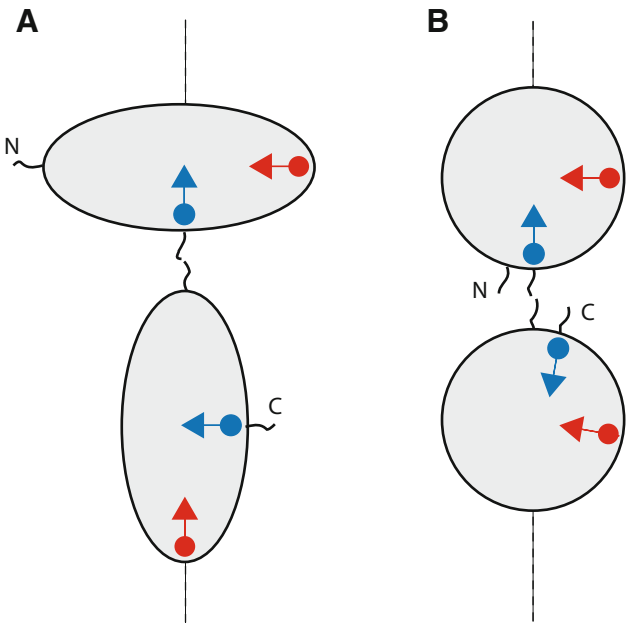

Fig. 4 Schematic structure of tandem domains with resident spins. *Blue circles* represent ^{15}N spins belonging to *i*-th residue in the N- and C-terminal domains (assuming that residue numbering is the same in the two domains); *blue arrows* represent the corresponding ^{15}N-^{1}H^{N} vectors. *Red circles/arrows* refer to the *j*-th residue. The *dashed line* corresponds to the long axis of the steric alignment tensor

properties of the two-domain constructs. The radii of gyration of the disintegrin and NBD domains are 8.8 and 15.3 Å, respectively. Let us accept that both domains are spherical. The average end-to-end distance in the random peptide chain, such as the computer-generated linker, scales with the number of residues as $N^{0.6}$ (Fitzkee and Rose 2004). From here we can estimate the 'scaling factor' that relates the NBD tandem to disintegrin tandem. For example, the NBD tandem with 25-residue linker can be viewed as a 'resized' copy of the disintegrin tandem with 10-residue linker. This is roughly consistent with the trends seen in Fig. 3a.

Generally speaking, N- and C-terminal domains in the tandem are inequivalent: one is attached to the linker through its C-terminal residues while the other through the N-terminal residues. In principle, this should lead to doubling of the spectral resonances. This situation is illustrated in Fig. 4a, which shows two domains connected in tandem. The scheme Fig. 4a suggests that some of the amide signals from N- and C-domains should be resolved with respect to their chemical shift. Furthermore, most of the signals should be distinguishable with respect to dipolar couplings. For instance, the equivalent residues from N- and C-terminal domains (blue circles) sense somewhat different chemical environment: one is located next to the fused terminus, while the other is close to the free terminus. Hence one can expect to observe two distinct HSQC peaks from these sites. Moreover, the orientation of the ^{15}N-^{1}H^{N} vectors at the two sites (blue arrows) relative to the long axis of the alignment tensor (dashed line) is also different. Hence the residual dipolar couplings for these two sites should be different as well. In fact, it is anticipated that most of the N- and C-terminal residues can be differentiated on the basis of their dipolar couplings. For instance, another pair of equivalent residues (red circles) illustrates the case where chemical shifts are likely to be degenerate, but RDCs are very different. Figure 4b represents a 'softer' version of the same scenario. When N- and C-termini of each domain are positioned close to each other, the RDCs from the two domains may turn out to be similar. Apparently, this scenario holds true for the α-spc SH3 tandem (see Experimental Results). Similar results were recently reported in the experimental study of GB1 construct featuring a short 3-residue linker (Walsh et al. 2010). Of the simulated constructs, see Fig. 3, NBD tandem illustrates the situation where the couplings from the two domains are reasonably close. The other two constructs, however, predict significant differences between the two domains.

Results

RDC data

The spectra of the dSH3 tandems turn out to be very similar to the spectrum of the isolated dSH3. For example, Fig. 5 shows the spectrum of the two-domain construct with long disordered linker superimposed on the spectrum of the single domain (spectral peaks colored red and blue, respectively). The resonances from multiple glycine residues of the random-coil GG(GGSGG)$_5$GG sequence are all grouped in the area of the spectrum where they do not interfere with intra-domain peaks (outlined by the green contour). Most of the peaks from N- and C-terminal domains overlap. Furthermore, they overlap or fall very close to the peaks from the isolated dSH3 domain (cf. red and blue contours in Fig. 5), so that the assignments (Blanco et al. 1997) can be transferred in a straightforward fashion. The transfer of resonance assignment from a single domain to a multidomain assembly is presently a common practice (Tzakos et al. 2006; Gelis et al. 2007; Sprangers and Kay 2007; Heikkinen et al. 2009; Brosey et al. 2009). The lack of resonance shifts between the isolated domain and the tandem confirms the absence of specific domain–domain interactions (along the same lines, ^{15}N relaxation study has not found any evidence of protein aggregation in α-spc SH3 (Chevelkov et al. 2007)).

As discussed above, a small number of residues that are located close to the linker attachment point give rise to a pair of resolved peaks, corresponding to the sites in the

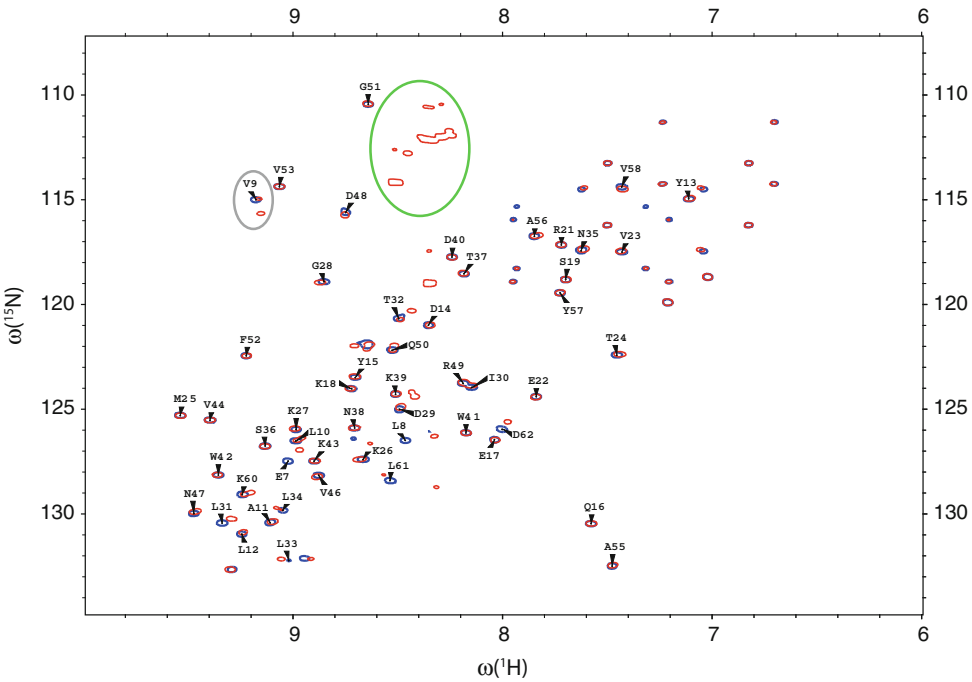

Fig. 5 HSQC spectra of the isolated dSH3 domain (*blue*) and dSH3-*ll*-dSH3 tandem (*red*). The dSH3 resonances are labeled except for Trp, Asn, and Gln side-chain correlations. The elliptical contours outline the resonances from the glycine-rich linker and from the residue V9 which demonstrates peak doubling effect

N- and C-terminal domains (one example of this behavior is residue V9, outlined by the grey contour in the spectral map Fig. 5). The effect is most pronounced in the short-linker construct, dSH3-*sl*-dSH3, where the two domains are spatially close. In this sample, 18 residues show the peak doubling effect, with half of them fully resolved. In particular, the peaks from terminal residues 7, 8 and 61, 62 are strongly affected. Inspection of the crystallographic structure shows that these two segments are connected: there is a backbone–backbone hydrogen bond between residues 8 and 61. Thus all of the four residues prove to be in the immediate vicinity of the linker region, resembling the arrangement shown in Fig. 4b. The split peaks associated with residues 7, 8, and 61, 62 show large chemical-shift separations, which makes it impossible to assign most of them. Furthermore, it is likely that the conformation of these residues in dSH3-*sl*-dSH3 differs from that seen in the crystallographic structure of fSH3. Considering these complications we choose to restrict the data analysis to the core region of the domains, residues 9–60. This restriction has been applied to all α-spc SH3 constructs investigated experimentally and via simulations.

A separate question arises as to whether the resonances from N- and C-terminal domains can be distinguished on the basis of residual dipolar couplings. In principle, one may expect that a pair of peaks that is completely overlapped in HSQC spectrum would become resolved in IPAP-HSQC (due to potentially large differences in RDCs, on the order of tens of Hz; see Fig. 4a and surrounding discussion). In reality, however, we have not found any examples of such behavior, despite high resolution of the IPAP-HSQC spectra. In the experiments involving dSH3-*sl*-dSH3, a few RDCs that have been measured separately for the sites in the N- and C-domains differ by no more than several Hz (e.g. in the case of V9 the difference is 0.5 Hz). At this level, RDCs alone cannot produce resolved spectral peaks. In summary, it appears that tandem dSH3 constructs behave as indicated in the scheme Fig. 4b: the N- and C-termini of each domain are aligned and together form a linker attachment point; the RDCs in the two domains are similar.

Given that only few residues allow for separate measurements of RDCs associated with N- and C-terminal domains and the measured values prove to be very similar, we choose to average such couplings and add the results to the bigger dataset consisting of the average RDCs. For instance, in the case of dSH3-*ll*-dSH3 we have measured 44 couplings that inherently represent the average values, plus two distinct couplings from the residue V9. After these two couplings are averaged, the final data set consists of 45 entries (pertaining to the region 9–60 in the dSH3 domain). The same approach has been applied to the other samples (Tables S2, S3).

The RDC data obtained as described above have been fitted to the crystallographic coordinates 1U06, chain A (see Materials and Methods). Although the experimental RDCs from the tandem constructs represent the averages between the N- and C-terminal domains, they have been fitted in the standard fashion, same as the data from the isolated dSH3 and fSH3 domains. As already commented, this approach is fully legitimate, so long as the internal structure of the two domains is identical. Indeed, residual

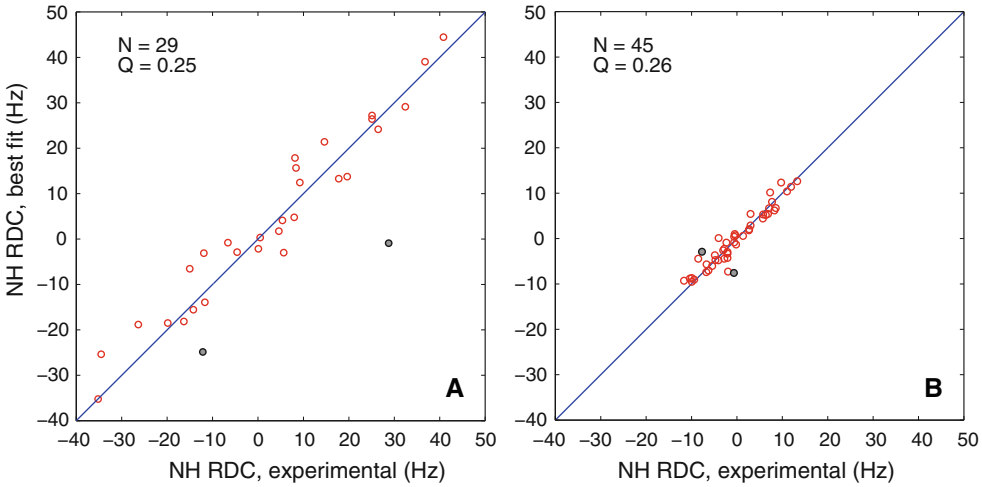

Fig. 6 Fits of the experimental RDCs from **a** the isolated dSH3 domain and **b** dSH3-*ll*-dSH3 tandem with the structural model based on the coordinate set 1U06 (protonated, energy minimized). The fitting is limited to the core portion of the protein, residues 9–60, and does not include the data from N47 and D48 (*grey circles* in the plot). The total number of the fitted RDCs and the figure of merit Q are indicated in the graphs

Table 2 Alignment parameters from a series of samples in $r = 0.96$ PEG/hexanol media

	dSH3	fSH3	dSH3-*sl*-dSH3	dSH3-*ml*-dSH3	dSH3-*ll*-dSH3
A_a (10^{-3})	2.23	−0.78	−1.50	−0.78	0.41
R	0.40	0.33	0.36	0.64	0.41
$\{\alpha, \beta, \gamma\}$	306°	32°	33°	36°	103°
	65°	39°	80°	83°	54°
	79°	321°	63°	64°	154°
N	43	48	41	40	45
Q	0.25	0.21	0.22	0.26	0.25
$\Delta(^2H)$ (Hz)	17.1	12.3	20.3	15.0	11.0
GDO (10^{-3})	2.76	1.32	1.54	1.18	0.80
NSP	1.0	0.90	−0.64	−0.47	0.21

The conventions are the same as in Table 1

dipolar couplings represent, by definition, the average between many molecular orientations. The quality of the RDC fits for dSH3 and dSH3-*ll*-dSH3 is illustrated in Fig. 6a and b, respectively. In both cases the agreement between the experimental and fitted couplings is good, as evidenced by the Q factors of 0.25 and 0.26 (typical of a small globular protein represented by a crystallographic structure with 1.5 Å resolution) (Bax 2003).

Two residues consistently show poor agreement in all fits—N47 and D48 (grey circles in Fig. 6). This discrepancy should be attributed to the local conformational variability at this particular site. In the crystal, N47 assumes an unusual conformation ($\phi \approx 50°, \psi \approx -110°$) (Vega et al. 2000) which has a significant destabilizing effect on the structure (Martinez et al. 1998). A similarly unfavorable conformation is seen in solution structure ($\phi \approx 60°, \psi \approx -75°$), except for one conformer (Blanco et al. 1997). The built-in strain apparently triggers local motions at this site. Both N47 and D48 display sharply elevated B-factors and several of the side-chain atoms in D48 cannot be resolved by x-ray crystallography at all. D48 is also the only residue featuring exchange broadening in solution ($R_{ex} = 28$ s^{-1} at 600 MHz) (Chevelkov et al. 2010). As it turns out, the RDC data from N47 and D48 can be accommodated using the existing structural models. Specifically, if these two residues are 'grafted' from the NMR structure 1AEY onto the crystallographic structure 1U06, some of the resulting hybrid models lead to a satisfactory RDC fit.

Alignment

The alignment parameters derived from multiple SH3 samples are summarized in Tables 1, 2 and in Fig. 7. As already discussed, special care was taken to ensure that the alignment conditions are the same throughout each series of samples. In the case of $r = 0.85$ samples, the residual quadrupolar splittings $\Delta(^2H)$ proved to be fairly uniform and reproducible (with the apparent exception of dSH3-*ml*-dSH3, see Table 1). For $r = 0.96$, the splittings were lower than expected and varied significantly from one sample to the other (see Table 2). This latter media appears to be only marginally stable at 25°C (Rückert and Otting 2000). Generally, the dilute PEG/hexanol media used in NMR experiments falls close to the boundary between lamellar

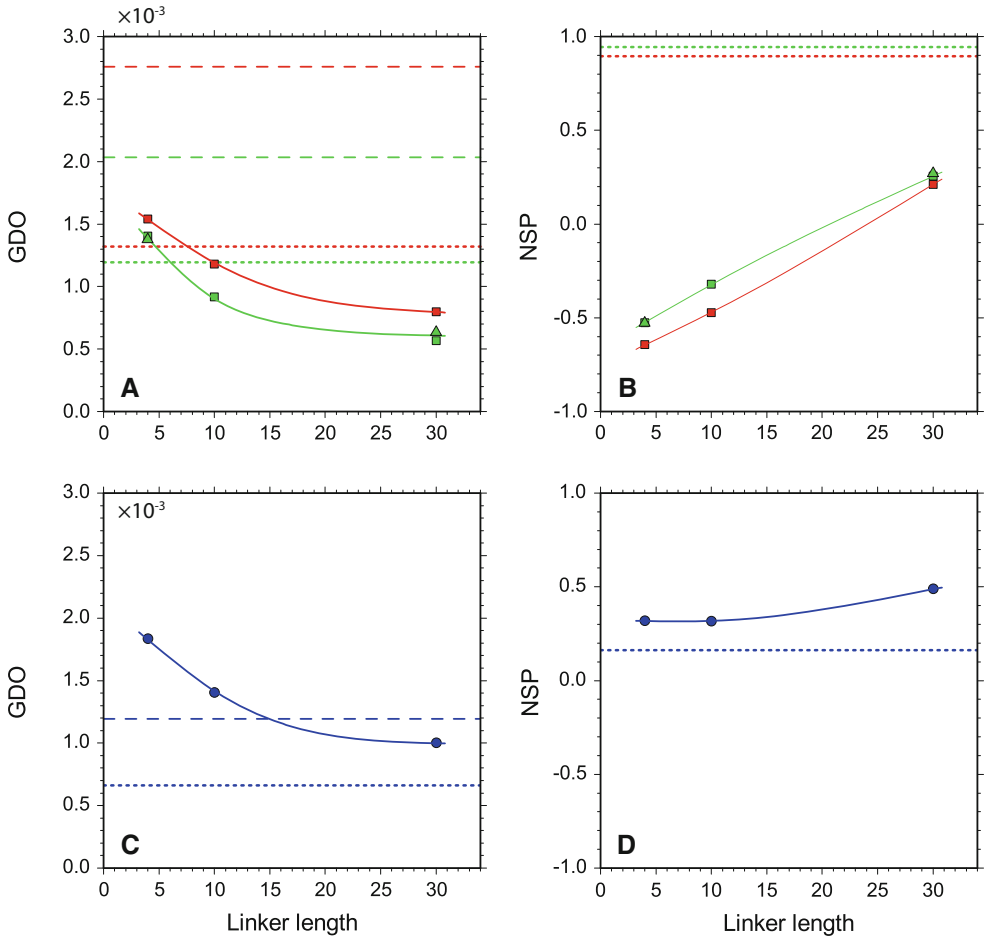

Fig. 7 GDO and NSP alignment parameters in a series of constructs of α-spc SH3 domain. **a** and **b** Experimental data from dSH3-*sl*-dSH3, dSH3-*ml*-dSH3, and dSH3-*ll*-dSH3 tandem constructs (*squares* for the main series of samples, *triangles* for replicate samples), as well as dSH3 and fSH3 single-domain constructs (*dashed* and *dotted horizontal lines*, respectively). The data from $r = 0.85$ and 0.96 samples are indicated in *green* and *red*, respectively. Spline curves are added to guide the eye. The NSP parameter is calculated relative to the dSH3 construct; the corresponding value for the dSH3 itself is 1.0. **c** and **d** The corresponding PALES-based simulations using the steric alignment model which assumes planar bicelle bilayers with the thickness 28.6 Å

phase L_α and a 'sponge phase' L_3, which itself has a complex and variable morphology (Gaemers and Bax 2001). In this situation, minor variations in the composition of the sample—possibly, the differences between the protein constructs themselves—may cause significant changes in the alignment. To account for these changes, we choose to normalize the GDO parameter according to the magnitude of $\Delta(^2H)$, see Table 1. In doing so we assume that $\Delta(^2H)$ reflects the aligning properties of the media not only with respect to water, but also to different α-spectrin SH3 constructs. This is an admittedly *ad hoc* approach, with only limited experimental justification (Rückert and Otting 2000). Note, however, that it appears justified for a series of samples where the alignment is dictated by the (invariant) properties of the α-spc SH3 domain. At least in the case of $r = 0.85$ media the alignment conditions seem to be highly reproducible from one sample to the other.

The consistency of different sample preparations can be verified through measurements on replicate samples. The results from the duplicate dSH3-*sl*-dSH3 and dSH3-*ll*-dSH3 samples are in very good agreement with each other (see Table 1; also compare squares and rectangles in Fig. 7). Furthermore, the data taken at $r = 0.85$ and 0.96 are also in good agreement insofar as rhombicity and the orientation of the alignment frame are concerned. One apparent exception is dSH3-*ml*-dSH3, cf. Tables 1 and 2. This construct displays the rhombicity R approaching the maximum value of 2/3. Under these circumstances, the labeling of the principal axes of the alignment tensor becomes ambiguous, causing apparent differences in the Euler angles. In fact, the two data sets are highly consistent, as confirmed by the NSP of the respective alignment tensors, 0.998.

The only sample where reproducibility truly appears to be an issue is dSH3. While the orientation of the alignment

frame is well reproduced, the rhombicity shows a significant amount of variation, Tables 1 and 2. Furthermore, the GDO value varies significantly from one preparation to the other. One thing is certain, however: the alignment of dSH3 is unexpectedly strong in all anisotropic samples, cf. Fig. 6 (also confirmed by the results from an additional dSH3 sample, not shown). In what follows, this observation is discussed in greater detail.

Figure 7a provides the experimental verification of the concept presented in this paper, namely, that the tandem with a short linker should align strongly, whereas the tandem with a long linker should align more weakly (given that the shape of the domains is close to spherical, the linker is unstructured, and there is no significant domain–domain interactions). Furthermore, it appears that the simple steric alignment model, as implemented in PALES, in conjunction with the conformational ensembles generated with the help of TraDES, provide an adequate explanation for the experimental observations (compare squares/triangles in Fig. 7a and circles in Fig. 7c). Note that the PALES-based prediction algorithm does not involve any adjustable parameters (see Materials & Methods). The NSP data in Fig. 7b confirm our hypothesis that with increase in linker length the tandem domains begin to behave as independent entities, i.e. align similarly to isolated domains.

From inspection of Fig. 7 it becomes obvious, however, that the steric alignment model is only partially successful in reproducing the experimental data. While the GDO values for tandem constructs are predicted with reasonable accuracy, the predictions for isolated domains, dSH3 and fSH3, are off the mark (cf. dashed/dotted lines in Fig. 7a, c). The case of the dSH3 is particularly instructive. This domain represents a spheroid with modest anisotropy, $I_{\parallel}/I_{\perp} = 0.83$. Its structure is well defined: only one N-terminal and one C-terminal residue are classified as 'unstructured'. Assuming that the mechanism of alignment is steric, dSH3 should show only a modest degree of alignment. Yet the experimental GDO value is exceedingly high (see also Fig. 6a). Moreover, the experimentally determined alignment frame bears no similarity to the predicted one, as indicated by the low NSP value, −0.11. This leads us to suggest that the alignment mechanism in the PEG/hexanol media is, in fact, different from the simple steric model postulated in PALES.

To further investigate this possibility, we turn to the literature reports of RDC measurements in the PEG/hexanol media. For ubiquitin, the orientation of the alignment frame is predicted by PALES with reasonable accuracy, NSP = 0.78–0.86. However, the GDO values are overestimated by a factor 2.1–4.5 (Lakomek et al. 2006; Briggman and Tolman 2003). Our own measurements on ubiquitin in PEG/hexanol media at pH 3.5 led to similar results, with GDO overestimated by 2.1–2.3 (data not shown). The agreement becomes even poorer if the results are scaled according to $\Delta(^2H)$. The data from another small protein, GB3 domain, produce NSP = 0.94 and another poor GDO prediction, off by a factor 3.6 (Ulmer et al. 2003). Taken together with our current data for α-spectrin SH3, these results suggest that PALES can both *under*- and *overestimate* the degree of alignment by a wide margin.

This shortcoming seems to be specific to the PEG/hexanol media. In the case of DMPC/DHPC bicelles, the steric PALES model fares much better. For ubiquitin, NSP = 0.96–0.97 and the predicted GDO value is only 1.3 times higher than the experimental value. For GB3 domain, NSP = 0.90 and the predicted GDO is essentially exact. Clearly, it would have been desirable to repeat our experiments in the DMPC/DHPC bicelles. Unfortunately, fSH3 and especially dSH3 constructs tend to precipitate at pH 6.5–7.0 required by this alignment media (van Rossum et al. 2003).

Alignment mechanisms

Unexpectedly, the results Fig. 7a indicate that nearly-spherical dSH3 domain aligns stronger than the markedly anisotropic dSH3-*sl*-dSH3 construct. In this section we discuss the potential sources of this behavior. While we cannot offer any single definitive explanation, it is useful to review various factors contributing to PEG/hexanol alignment.

The result in Fig. 7 appears paradoxical only in the light of the simple steric model involving planar bicelles. Alternatively, if the morphology of the media is complex, then a steric mechanism can, in principle, lead to an arbitrary alignment tensor. Indeed, assume for the sake of argument that anisotropic media contains cavities that can (transiently) accommodate certain proteins. Assume further that the cavities are shaped such that they form 'lock and key' pair with a given protein. Clearly, the resulting alignment tensor will be different from the one that is predicted by PALES.

In fact, recent studies showed that the topology of the most used liquid-crystal media is far more complex than previously believed. The DMPC/DHPC bicelles, which were originally described as large disks, have been instead visualized as "two-dimensional networks of flattened, highly branched, cylindrical micelles and lamellar sheets perforated by large irregular holes" (van Dam et al. 2006). Similar observations have been made with regard to the PEG-hexanol media. Early on it has been noted that the L_α bilayers display collective wave-like modulation (undulation) (Strey et al. 1990; Jonstromer and Strey 1992). Gaemers and Bax pointed out that dilute conditions used in

NMR experiments bring the system close to the transitions region where the lamellar phase L_α coexists with a 'sponge phase' L_3 (Gaemers and Bax 2001). Their diffusion data further suggest that the motion of probe molecules between the oriented bilayers (i.e. along the z-axis in the NMR tube) is significantly obstructed. These obstructions may arise from the 'bridges' between the adjacent bilayers, the presence of small globular structures (e.g. micelles) in the interbilayer space, etc. None of this potentially complex topology is taken into consideration in the PALES model—which may explain its apparent lack of accuracy, Fig. 7.

From a more general perspective, the failure of PALES can be attributed to site-specific interactions between the proteins and the media. Generally, PEG-based media are protein-friendly in a sense that there is little protein adsorption. Nevertheless, specific interactions involving PEG-hexanol have been observed for short peptides (Ohnishi and Shortle 2003), N-terminal domain of the enzyme I from bacterial phosphotransferase system (Suh et al. 2008), two cytoplasmic domains of chitobiose transporter (Jung et al. 2010), and folding intermediate of FF domain from HYPA/FBP11 (Korzhnev et al. 2010). There is no doubt that many more cases remain undocumented (dismissed as unsuccessful attempts to prepare aligned samples).

PEG doped with hexanol forms a bilayer where the surface is lined with hydroxyl headgroups. The interactions of the proteins with such bilayers have been modeled in several influential theoretical studies (Jeon et al. 1991; McPherson et al. 1998). All-atom force field simulations have been also reported (Lim and Herron 1992; Zheng et al. 2004; Zheng et al. 2005). These simulations clearly identify the preferred orientation of the protein (lysozyme) on a surface of a PEG bilayer. Of note, the binding interface appears to be formed by charged/polar residues interacting with PEG oxygens at or near the surface of the bilayer (Zheng et al. 2004). Of course, the binding is only weakly specific since many sites on the protein surface can form such favorable contacts. Note that electrostatic effects play a prominent role in molecular alignment even when the media is not charged—it is sufficient that the media is polar. These aspects has been extensively discussed in the small-molecule studies (Emsley et al. 1991; Terzis and Photinos 1994; Syvitski and Burnell 2000; Dingemans et al. 2003; Wardle et al. 2004).

To elucidate a possible role of electrostatics in PEG/hexanol media, we have prepared an additional sample of fSH3 containing 100 mM NaCl. If the alignment is purely steric, as is commonly held, then one would not expect to observe any significant changes upon addition of NaCl. In reality, however, the GDO parameter jumps from 1.19×10^{-3} to 1.85×10^{-3}, see Table 1.[1] This result is unexpected and, at a first glance, counterintuitive. Indeed, if electrostatic interactions are involved, then addition of salt should screen out these interactions and thus *lower* the degree of alignment. In fact, a number of plausible explanations can be suggested for this finding. The increase in GDO may result from the interplay between the electrostatic and steric mechanisms. Furthermore, it may also occur in the context of purely electrostatic alignment. As described by Wu and co-workers (Wu et al. 2006), excessive electrostatic repulsion drives the solute away from the charged media and into the bulk, where it cannot effectively align. Adding NaCl to the solution reverses this process, thus increasing the degree of alignment. Details of alignment notwithstanding, the observed dependence of the GDO on the ionic strength of the solution suggests that an electrostatic mechanism plays a role.

To further probe the relevance of electrostatic interactions, we have undertaken an additional series of PALES simulations. Unlike the steric model, which has no adjustable parameters, the electrostatic model contains several such quantities. As described in the Materials & Methods section, the ionic strength of the solution was set to the actual experimental value, 17 mM. The other two parameters, pH and the surface charge density of the medium σ, were optimized in an *ad hoc* fashion to obtain the best possible agreement with the five experimental GDO values (listed in Table 1). The optimization was conducted by means of the grid-search in the space of two parameters. Calculations at each point on the grid involved multiple PALES jobs (addressing five conformational ensembles, each comprised of several thousand conformers).

The optimized electrostatic model produced the predictions for GDO and NSP as shown in Fig. 8. The main result of this simulation is that it is indeed possible to recreate the situation where dSH3 and fSH3 display a higher degree of alignment than tandem constructs, as observed experimentally (see Fig. 7a). The parameters obtained from the model optimization are $\sigma = 0.05$ e/nm^2, effective pH 4.8. The value 0.05 is relatively low in comparison, for example, with Pf1 phage, -0.47 e/nm^2. In principle, surface charge may arise from adsorption of ions on the interface between water and non-ionic surfactants (Elworthy et al. 1971; Balzer 1993; Karraker and Radke 2002; Stubenrauch and von Klitzing 2003; Wardle et al. 2004). However, it appears that the charge density on the surface of a C12E5 bilayer in water is actually much lower than 0.05 e/nm^2 and negative (Balzer 1993; Stubenrauch and von Klitzing 2003), even in the presence of a small amount of ionic surfactant (Schomacker and Strey 1994; von Berlepsch and de Vries 2000). At the same time, the obtained effective pH value, 4.8 is significantly higher than the actual experimental pH, 3.5.

[1] In contrast, we found that alignment of ubiquitin in PEG/hexanol is insensitive to salt.

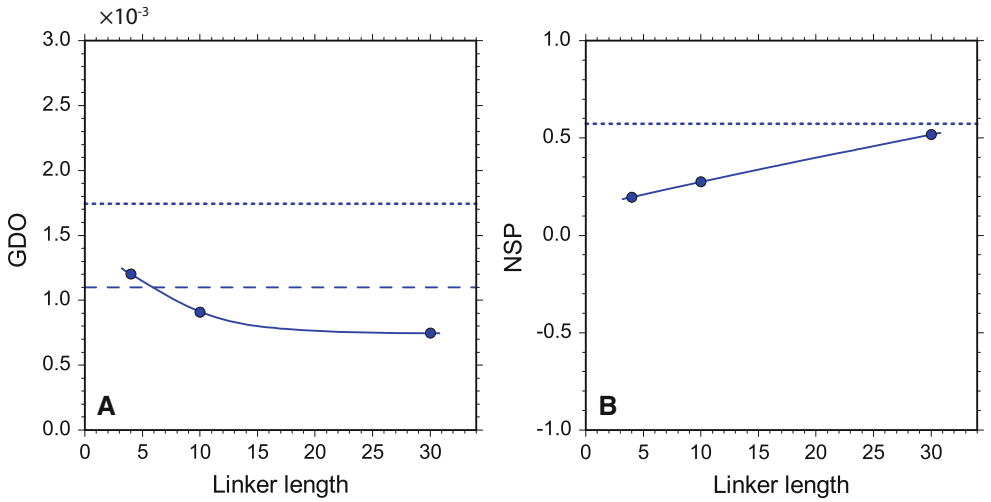

Fig. 8 The GDO and NSP alignment parameters for α-spc SH3 constructs as obtained from PALES-based electrostatic simulations. The optimized model assumes planar bicelle bilayers with surface charge density $\sigma = 0.05$ e/nm^2, and effective sample pH 4.8

In judging the results of these simulations one should keep in mind that our model of electrostatic alignment is rather crude. Specifically, the pK_a values used in the calculations have a typical error of 0.5–1 unit, with maximum error of up to 2 units (Li et al. 2005). Furthermore, the predictions of electrostatic alignment are highly sensitive to fine details of the charge distribution. For instance, the predicted alignment parameters may change in response to a conformational jump of a single charged side chain. To demonstrate this effect, we regenerated the SH3 conformational ensembles using the BAAAA structure of the SH3 domain, i.e. a structure where the single side chain, E17, is placed in an alternative conformation (see Materials & Methods). This minor structural modification led to significant decrease in the predicted GDO of the fSH3 sample, from 1.74×10^{-3} to 1.36×10^{-3}.

In conclusion, the presented PALES-based electrostatic simulation demonstrates the feasibility of the scenario in which dSH3 and fSH3 align more efficiently than the tandem constructs. Beyond that, we do not attach any particular significance to the results shown in Fig. 8. One can envisage that more sophisticated prediction tools will be developed in future, where a fully atomistic approach will be extended to the liquid crystal media. It can also be expected that MD-based methods will eventually emerge. The examples of such developments have been cited above (Dingemans et al. 2003; Wardle et al. 2004; Zangi and Engberts 2005; Zheng et al. 2004; Polyansky et al. 2005; Zheng et al. 2005; Aroulanda et al. 2006). The progress in this area, however, is limited by our relatively poor knowledge of the liquid crystal structure and dynamics.

Conclusion

In this study we focus on proteins comprised of two modular domains connected through a random-coil linker (α-spc SH3 domains connected through glycine-serine linker). When the linker is short, the tandem SH3 construct represents a dumbbell-shaped molecule with relatively little domain–domain mobility. Being placed in steric alignment media, such as PEG/hexanol, this molecule is expected to display a high degree of alignment. As the length of the linker is increased, the domains become effectively uncoupled and start behaving as independent entities. Consequently, the degree of alignment observed in each of the tandem domains approaches that of the isolated SH3 domain.

In qualitative terms, it has been noted before that differential domain alignment conveys information about interdomain mobility (Braddock et al. 2001; Ulmer et al. 2002; Poon et al. 2007; Holland et al. 2008; Wang et al. 2009). Here we attempted to quantify the dependence of the alignment parameters (amplitude and orientation) on the length of the linker. The form of the corresponding transition curves depends on the details of the system, e.g. the shape of the individual domains and the linker attachment points. In the case of electrostatic alignment, the trend can be easily reversed, i.e. the degree of alignment may increase with lengthening of the linker. The transition phenomenon itself, however, is sufficiently general.

To translate the proposed intuitive model into quantitative predictions, we generated a series of conformational ensembles representing α-spc SH3 constructs with different linker lengths (the disordered linkers and terminal sequences

were built using the program TraDES). These ensembles were used to predict the parameters of steric alignment (program PALES). To test the validity of the theoretical predictions, we prepared three tandem constructs of α-spc SH3 and two variants of the isolated SH3 domain. These chimeras were used to measure backbone $^{1}H^{N}$-^{15}N RDCs in two series of PEG/hexanol samples ($r = 0.85$ and 0.96). It was found that the alignment of tandem proteins is in line with expectations and agrees reasonably well with the results from PALES-based simulations. The isolated domains, however, displayed much stronger alignment than expected. Furthermore, the degree of alignment turned out to be sensitive to the ionic strength of the solution, which is generally not expected to be the case for a well-behaved steric alignment media.

Several factors may potentially contribute to this situation. First, the topology of the liquid crystal media is known to be more complex than envisaged in the simple PALES model (i.e. planar bilayers). Furthermore, there is a possibility that the phase equilibrium of the liquid crystal changes in response to addition of protein or salt (these changes may not necessarily be fully reflected in $\Delta(^{2}H)$) (Ottiger and Bax 1998). It is also likely that protein alignment in PEG/hexanol significantly depends on (weak) site-specific interactions. In particular, an electrostatic mechanism appears to play a role, visibly affecting the alignment of isolated domains.

In summary, it seems that PEG/hexanol bears little resemblance to an idealized alignment media envisioned in PALES. One would be well-advised to keep this in mind, especially in the context of studies where PALES is used to simulate RDC data from disordered proteins dissolved in PEG-based media (Alexandrescu and Kammerer 2003; Bernado et al. 2005; Jensen et al. 2008; Pinheiro et al. 2011). Indeed, if we cannot reliably predict the alignment of globular proteins, the predictions should be even less reliable for disordered proteins. In this latter case, all residues can potentially form site-specific contacts with the media and, furthermore, these contacts may significantly skew the original conformational equilibrium (for those protein molecules that are in the vicinity of the media and give rise to observable RDCs).

The chimera protein investigated in this work is comprised of two non-interacting domains connected by a variable-length random linker. It can thus be described as an example of the system with *minimal domain coupling*. This example provides an important point of reference for future studies of the proteins with substantial domain coupling, either through domain–domain interactions (e.g. Crk adaptor protein (Kobashigawa et al. 2007), Syk kinase (Zhang et al. 2008)) or through a structured linker (e.g. calmodulin (Ikura et al. 1991), troponin (Slupsky et al. 1995)). In the case of popular PEG/hexanol media, the usefulness of this approach is presently limited by lack of precise information about the morphology of the oriented phase and the details of alignment mechanism. We expect that the DMPC/DHPC media offers a better chance of success, although it is far more restrictive in terms of sample conditions. Further progress in this area depends on development of more accurate alignment prediction tools. Ultimately, such tools should be based on all-atom MD simulations modeling the interactions between the protein and the alignment media.

Acknowledgments This work has been funded through NSF grant MCB-044563. We thank Ryan Muir for his help with early versions of the structure-generating scripts. We are also grateful to Etti Harms, Nina Gorenstein, Josh Ward, and Yi Xue for their advice on different aspects of experimental work.

References

Akakura S, Kar B, Singh S, Cho L, Tibrewal N, Sanokawa-Akakura R, Reichman C, Ravichandran KS, Birge RB (2005) C-terminal SH3 domain of Crkll regulates the assembly and function of the DOCK180/ELMO Rac-GEF. J Cell Physiol 204:344–351

Alexandrescu AT, Kammerer RA (2003) Structure and disorder in the ribonuclease S-peptide probed by NMR residual dipolar couplings. Protein Sci 12:2132–2140

Andre I, Linse S, Mulder FAA (2007) Residue-specific pK_a determination of lysine and arginine side chains by indirect ^{15}N and ^{13}C NMR spectroscopy: Application to apo calmodulin. J Am Chem Soc 129:15805–15813

Aroulanda C, Celebre G, De Luca G, Longeri M (2006) Molecular ordering and structure of quasi-spherical solutes by liquid crystal NMR and Monte Carlo simulations: the case of norbornadiene. J Phys Chem B 110:10485–10496

Bae SH, Dyson HJ, Wright PE (2009) Prediction of the rotational tumbling time for proteins with disordered segments. J Am Chem Soc 131:6814–6821

Balzer D (1993) Cloud point phenomena in the phase behavior of alkyl polyglucosides in water. Langmuir 9:3375–3384

Bas DC, Rogers DM, Jensen JH (2008) Very fast prediction and rationalization of pK_a values for protein-ligand complexes. Proteins 73:765–783

Bax A (2003) Weak alignment offers new NMR opportunities to study protein structure and dynamics. Protein Sci 12:1–16

Bax A, Tjandra N (1997) High-resolution heteronuclear NMR of human ubiquitin in an aqueous liquid crystalline medium. J Biomol NMR 10:289–292

Bax A, Kontaxis G, Tjandra N (2001) Dipolar couplings in macromolecular structure determination. Method Enzymol 339:127–174

Berlin K, O'Leary DP, Fushman D (2009) Improvement and analysis of computational methods for prediction of residual dipolar couplings. J Magn Reson 201:25–33

Bernado P, Blanchard L, Timmins P, Marion D, Ruigrok RWH, Blackledge M (2005) A structural model for unfolded proteins from residual dipolar couplings and small-angle x-ray scattering. Proc Natl Acad Sci USA 102:17002–17007

Bertini I, Del Bianco C, Gelis I, Katsaros N, Luchinat C, Parigi G, Peana M, Provenzani A, Zoroddu MA (2004) Experimentally exploring the conformational space sampled by domain reorientation in calmodulin. Proc Natl Acad Sci USA 101:6841–6846

Bertini I, Gupta YK, Luchinat C, Parigi G, Peana M, Sgheri L, Yuan J (2007) Paramagnetism-based NMR restraints provide maximum allowed probabilities for the different conformations of partially independent protein domains. J Am Chem Soc 129:12786–12794

Bewley CA (2001) Solution structure of a cyanovirin-N: Manα1-2Manα complex: structural basis for high-affinity carbohydrate-mediated binding to gp120. Structure 9:931–940

Blanco FJ, Ortiz AR, Serrano L (1997) ^{1}H and ^{15}N NMR assignment and solution structure of the SH3 domain of spectrin: comparison of unrefined and refined structure sets with the crystal structure. J Biomol NMR 9:347–357

Braddock DT, Cai ML, Baber JL, Huang Y, Clore GM (2001) Rapid identification of medium- to large-scale interdomain motion in modular proteins using dipolar couplings. J Am Chem Soc 123:8634–8635

Braddock DT, Baber JL, Levens D, Clore GM (2002) Molecular basis of sequence-specific single-stranded DNA recognition by KH domains: solution structure of a complex between hnRNP KKH3 and single-stranded DNA. EMBO J 21:3476–3485

Briggman KB, Tolman JR (2003) De Novo determination of bond orientations and order parameters from residual dipolar couplings with high accuracy. J Am Chem Soc 125:10164–10165

Brooks BR, Brooks CL, Mackerell AD, Nilsson L, Petrella RJ, Roux B, Won Y, Archontis G, Bartels C, Boresch S, Caflisch A, Caves L, Cui Q, Dinner AR, Feig M, Fischer S, Gao J, Hodoscek M, Im W, Kuczera K, Lazaridis T, Ma J, Ovchinnikov V, Paci E, Pastor RW, Post CB, Pu JZ, Schaefer M, Tidor B, Venable RM, Woodcock HL, Wu X, Yang W, York DM, Karplus M (2009) CHARMM: the biomolecular simulation program. J Comput Chem 30:1545–1614

Brosey CA, Chagot ME, Ehrhardt M, Pretto DI, Weiner BE, Chazin WJ (2009) NMR analysis of the architecture and functional remodeling of a modular multidomain protein, RPA. J Am Chem Soc 131:6346–6347

Bryngelson JD, Wolynes PG (1987) Spin glasses and the statistical mechanics of protein folding. Proc Natl Acad Sci USA 84: 7524–7528

Bryngelson JD, Onuchic JN, Socci ND, Wolynes PG (1995) Funnels, pathways, and the energy landscape of protein folding: a synthesis. Proteins Struct Funct Genet 21:167–195

Cesareni G, Gimona M, Sudol M, Yaffe M (eds) (2005) Modular protein domains. Wiley-VCH, Weinheim

Chen K, Tjandra N (2008) Extended model free approach to analyze correlation functions of multidomain proteins in the presence of motional coupling. J Am Chem Soc 130:12745–12751

Chevelkov V, Faelber K, Diehl A, Heinemann U, Oschkinat H, Reif B (2005) Detection of dynamic water molecules in a microcrystalline sample of the SH3 domain of α-spectrin by MAS solid-state NMR. J Biomol NMR 31:295–310

Chevelkov V, Zhuravleva AV, Xue Y, Reif B, Skrynnikov NR (2007) Combined analysis of ^{15}N relaxation data from solid- and solution-state NMR spectroscopy. J Am Chem Soc 129: 12594–12595

Chevelkov V, Xue Y, Linser R, Skrynnikov NR, Reif B (2010) Comparison of solid-state dipolar couplings and solution relaxation data provides insight into protein backbone dynamics. J Am Chem Soc 132:5015–5017

Cipres A, Abassi YA, Vuori K (2007) Abl functions as a negative regulator of Met-induced cell motility via phosphorylation of the adapter protein CrkII. Cell Signal 19:1662–1670

Cornilescu G, Marquardt JL, Ottiger M, Bax A (1998) Validation of protein structure from anisotropic carbonyl chemical shifts in a dilute liquid crystalline phase. J Am Chem Soc 120:6836–6837

de la Torre JG, Huertas ML, Carrasco B (2000) Calculation of hydrodynamic properties of globular proteins from their atomic-level structure. Biophys J 78:719–730

Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A (1995) NMRPipe—a multidimensional spectral processing system based on unix pipes. J Biomol NMR 6:277–293

Denisov AY, Kloser E, Gray DG, Mittermaier AK (2010) Protein alignment using cellulose nanocrystals: practical considerations and range of application. J Biomol NMR 47:195–204

Dingemans T, Photinos DJ, Samulski ET, Terzis AF, Wutz C (2003) Ordering of apolar and polar solutes in nematic solvents. J Chem Phys 118:7046–7061

Dominy BN, Brooks CL (1999) Development of a generalized Born model parametrization for proteins and nucleic acids. J Phys Chem B 103:3765–3773

Donaldson LW, Gish G, Pawson T, Kay LE, Forman-Kay JD (2002) Structure of a regulatory complex involving the Abl SH3 domain, the Crk SH2 domain, and a Crk-derived phosphopeptide. Proc Natl Acad Sci USA 99:14053–14058

Elworthy PH, Rogers JA, Florence AT (1971) Stabilization of oil-in-water emulsions by nonionic detergents. 5. Effect of salts on rates of coalescence in a chlorobenzene emulsion. J Colloid Interf Sci 35:23–33

Emsley JW, Palke WE, Shilstone GN (1991) The inclusion of electrostatic and dispersion interactions into potentials of mean torque for solutes dissolved in uniaxial liquid-crystal solvents. Liq Cryst 9:643–648

Feldman HJ, Hogue CWV (2000) A fast method to sample real protein conformational space. Proteins Struct Funct Genet 39:112–131

Feller SM (2001) Crk family adaptors—signalling complex formation and biological roles. Oncogene 20:6348–6371

Fischer D, Geyer A (2005) NMR spectroscopic characterization of the membrane affinity of polyols. Magn Reson Chem 43:893–901

Fitzkee NC, Rose GD (2004) Reassessing random-coil statistics in unfolded proteins. Proc Natl Acad Sci USA 101:12497–12502

Freyssingeas É, Nallet F, Roux D (1996) Measurement of the membrane flexibility in lamellar and ''sponge'' phases of the $C_{12}E_5$/hexanol/water system. Langmuir 12:6028–6035

Futterer K, Wong J, Grucza RA, Chan AC, Waksman G (1998) Structural basis for syk tyrosine kinase ubiquity in signal transduction pathways revealed by the crystal structure of its regulatory SH2 domains bound to a dually phosphorylated ITAM peptide. J Mol Biol 281:523–537

Gaemers S, Bax A (2001) Morphology of three lyotropic liquid crystalline biological NMR media studied by translational diffusion anisotropy. J Am Chem Soc 123:12343–12352

Gelis I, Bonvin AMJJ, Keramisanou D, Koukaki M, Gouridis G, Karamanou S, Economou A, Kalodimos CG (2007) Structural basis for signal-sequence recognition by the translocase motor SecA as determined by NMR. Cell 131:756–769

Gordon JC, Myers JB, Folta T, Shoja V, Heath LS, Onufriev A (2005) H++: a server for estimating pK_as and adding missing hydrogens to macromolecules. Nucl Acids Res 33:W368–W371

Hakansson KO (2009) The structure of Mg-ATPase nucleotide-binding domain at 1.6 angstrom resolution reveals a unique ATP-binding motif. Acta Crystallogr D 65:1181–1186

Heikkinen O, Permi P, Koskela H, Ylanne J, Kilpelainen I (2009) ^{1}H, ^{13}C and ^{15}N resonance assignments of the human filamin A tandem immunoglobulin-like domains 16–17 and 18–19. Biomol NMR Assign 3:53–56

Holland NB, Nishimiya Y, Tsuda S, Sonnichsen FD (2008) Two domains of RD3 antifreeze protein diffuse independently. Biochemistry 47:5935–5941

Huston JS, Levinson D, Mudgetthunter M, Tai MS, Novotny J, Margolies MN, Ridge RJ, Bruccoleri RE, Haber E, Crea R, Oppermann H (1988) Protein engineering of antibody-binding sites: recovery of specific activity in an anti-digoxin single-chain

Fv analog produced in *Escherichia Coli*. Proc Natl Acad Sci USA 85:5879–5883

Ikegami T, Verdier L, Sakhaii P, Grimme S, Pescatore B, Saxena K, Fiebig KM, Griesinger C (2004) Novel techniques for weak alignment of proteins in solution using chemical tags coordinating lanthanide ions. J Biomol NMR 29:339–349

Ikura M, Kay LE, Krinks M, Bax A (1991) Triple-resonance multidimensional NMR study of calmodulin complexed with the binding domain of skeletal muscle Myosin Light-Chain Kinase: indication of a conformational change in the central helix. Biochemistry 30:5498–5504

Jensen MR, Houben K, Lescop E, Blanchard L, Ruigrok RWH, Blackledge M (2008) Quantitative conformational analysis of partially folded proteins from residual dipolar couplings: Application to the molecular recognition element of Sendai virus nucleoprotein. J Am Chem Soc 130:8055–8061

Jeon SI, Lee JH, Andrade JD, De Gennes PG (1991) Protein surface interactions in the presence of polyethylene oxide. I. Simplified theory. J Colloid Interf Sci 142:149–158

Jonstromer M, Strey R (1992) Nonionic bilayers in dilute solutions: effect of additives. J Phys Chem 96:5993–6000

Jung YS, Cai ML, Clore GM (2010) Solution structure of the $IIA^{Chitobiose}$-$IIB^{Chitobiose}$ complex of the N,N'-diacetylchitobiose branch of the *Escherichia Coli* phosphotransferase system. J Biol Chem 285:4173–4184

Karraker KA, Radke CJ (2002) Disjoining pressures zeta potentials and surface tensions of aqueous non-ionic surfactant/electrolyte solutions: theory and comparison to experiment. Adv Colloid Interfac 96:231–264

Keizers PHJ, Desreux JF, Overhand M, Ubbink M (2007) Increased paramagnetic effect of a lanthanide protein probe by two-point attachment. J Am Chem Soc 129:9292–9293

Khandogin J, Brooks CL (2006) Toward the accurate first-principles prediction of ionization equilibria in proteins. Biochemistry 45:9363–9373

Kobashigawa Y, Sakai M, Naito M, Yokochi M, Kumeta H, Makino Y, Ogura K, Tanaka S, Inagaki F (2007) Structural basis for the transforming activity of human cancer-related signaling adaptor protein CRK. Nat Struct Mol Biol 14:503–510

Koenig BW, Hu JS, Ottiger M, Bose S, Hendler RW, Bax A (1999) NMR measurement of dipolar couplings in proteins aligned by transient binding to purple membrane fragments. J Am Chem Soc 121:1385–1386

Korzhnev DM, Religa TL, Banachewicz W, Fersht AR, Kay LE (2010) A transient and low-populated protein-folding intermediate at atomic resolution. Science 329:1312–1316

Lakomek NA, Carlomagno T, Becker S, Griesinger C, Meiler J (2006) A thorough dynamic interpretation of residual dipolar couplings in ubiquitin. J Biomol NMR 34:101–115

Lazar GA, Desjarlais JR, Handel TM (1997) De novo design of the hydrophobic core of ubiquitin. Protein Sci 6:1167–1178

Li H, Robertson AD, Jensen JH (2005) Very fast empirical prediction and rationalization of protein pK_a values. Proteins 61:704–721

Lim K, Herron JN (1992) Molecular simulation of protein-PEG interactions. In Harris JM (ed) Poly(ethylene glycol) chemistry: biotechnical and biomedical applications. Plenum Press, New York, pp 29–56

Linghu H, Tsuda M, Makino Y, Sakai M, Watanabe T, Ichihara S, Sawa H, Nagashima K, Mochizuki N, Tanaka S (2006) Involvement of adaptor protein Crk in malignant feature of human ovarian cancer cell line MCAS. Oncogene 25:3547–3556

Maltsev AS, Ahmed AH, Fenwick MK, Jane DE, Oswald RE (2008) Mechanism of partial agonism at the GluR2 AMPA receptor: Measurements of lobe orientation in solution. Biochemistry 47:10600–10610

Martinez JC, Pisabarro MT, Serrano L (1998) Obligatory steps in protein folding and the conformational diversity of the transition state. Nat Struct Biol 5:721–729

McPherson T, Kidane A, Szleifer I, Park K (1998) Prevention of protein adsorption by tethered poly(ethylene oxide) layers: experiments and single-chain mean-field analysis. Langmuir 14:176–186

Miller CT, Chen G, Gharib TG, Wang H, Thomas DG, Misek DE, Giordano TJ, Yee J, Orringer MB, Hanash SM, Beer DG (2003) Increased c-Crk proto-oncogene expression is associated with an aggressive phenotype in lung adenocarcinomas. Oncogene 22:7950–7957

Mohana-Borges R, Goto NK, Kroon GJA, Dyson HJ, Wright PE (2004) Structural characterization of unfolded states of apomyoglobin using residual dipolar couplings. J Mol Biol 340: 1131–1142

Moltke S, Grzesiek S (1999) Structural constraints from residual tensorial couplings in high resolution NMR without an explicit term for the alignment tensor. J Biomol NMR 15:77–82

Moreno-Murciano MP, Monleon D, Marcinkiewcz C, Calvete JJ, Celda B (2003) NMR solution structure of the non-RGD disintegrin obtustatin. J Mol Biol 329:135–145

Muralidharan V, Dutta K, Cho J, Vila-Perello M, Raleigh DP, Cowburn D, Muir TW (2006) Solution structure and folding characteristics of the C-terminal SH3 domain of c-Crk-II. Biochemistry 45:8874–8884

Musacchio A, Noble M, Pauptit R, Wierenga R, Saraste M (1992) Crystal structure of a Src-homology 3 (SH3) domain. Nature 359:851–855

Nishihara H, Tanaka S, Tsuda M, Oikawa S, Maeda M, Shimizu M, Shinomiya H, Tanigami A, Sawa H, Nagashima K (2002) Molecular and immunohistochemical analysis of signaling adaptor protein Crk in human cancers. Cancer Lett 180:55–61

Ogawa S, Toyoshima H, Kozutsumi H, Hagiwara K, Sakai R, Tanaka T, Hirano N, Mano H, Yazaki Y, Hirai H (1994) The C-Terminal SH3 domain of the mouse c-Crk protein negatively regulates tyrosine phosphorylation of Crk associated p130 in rat 3Y1 cells. Oncogene 9:1669–1678

Ohnishi S, Shortle D (2003) Observation of residual dipolar couplings in short peptides. Proteins Struct Funct Genet 50:546–551

Opella SJ, De Angelis AA (2007) Bicelle samples for solid-state NMR of membrane proteins. Nat Protoc 2:2332–2338

Ottiger M, Bax A (1998) Characterization of magnetically oriented phospholipid micelles for measurement of dipolar couplings in macromolecules. J Biomol NMR 12:361–372

Ottiger M, Delaglio F, Bax A (1998) Measurement of *J* and dipolar couplings from simplified two-dimensional NMR spectra. J Magn Reson 131:373–378

Peterson ME, Long EO (2008) Inhibitory receptor signaling via tyrosine phosphorylation of the adaptor Crk. Immunity 29: 578–588

Pinheiro AS, Marsh JA, Forman-Kay JD, Peti W (2011) Structural Signature of the MYPT1-PP1 Interaction. J Am Chem Soc 133:73–80

Polyansky AA, Volynsky PE, Nolde DE, Arseniev AS, Efremov RG (2005) Role of lipid charge in organization of water/lipid bilayer interface: Insights via computer simulations. J Phys Chem B 109:15052–15059

Poon DKY, Withers SG, McIntosh LP (2007) Direct demonstration of the flexibility of the glycosylated proline-threonine linker in the Cellulomonas fimi xylanase Cex through NMR spectroscopic analysis. J Biol Chem 282:2091–2100

Prestegard JH, Al-Hashimi HM, Tolman JR (2000) NMR structures of biomolecules using field oriented media and residual dipolar couplings. Q Rev Biophys 33:371–424

Rodriguez-Castaneda F, Haberz P, Leonov A, Griesinger C (2006) Paramagnetic tagging of diamagnetic proteins for solution NMR. Magn Reson Chem 44:S10–S16

Rückert M, Otting G (2000) Alignment of biological macromolecules in novel nonionic liquid crystalline media for NMR experiments. J Am Chem Soc 122:7793–7797

Sarkar P, Reichman C, Saleh T, Birge RB, Kalodimos CG (2007) Proline cis-trans isomerization controls autoinhibition of a signaling protein. Mol Cell 25:413–426

Sass J, Cordier F, Hoffmann A, Rogowski M, Cousin A, Omichinski JG, Lowen H, Grzesiek S (1999) Purple membrane induced alignment of biological macromolecules in the magnetic field. J Am Chem Soc 121:2047–2055

Sass HJ, Musco G, Stahl SJ, Wingfield PT, Grzesiek S (2000) Solution NMR of proteins within polyacrylamide gels: Diffusional properties and residual alignment by mechanical stress or embedding of oriented purple membranes. J Biomol NMR 18:303–309

Schomacker R, Strey R (1994) Effect of ionic surfactants on nonionic bilayers: bending elasticity of weakly charged membranes. J Phys Chem 98:3908–3912

Schwieters CD, Clore GM (2001) Internal coordinates for molecular dynamics and minimization in structure determination and refinement. J Magn Reson 152:288–302

Schwieters CD, Kuszewski JJ, Tjandra N, Clore GM (2003) The Xplor-NIH NMR molecular structure determination package. J Magn Reson 160:65–73

Shakhnovich EI, Gutin AM (1989) Formation of unique structure in polypeptide chains. Theoretical investigation with the aid of a replica approach. Biophys Chem 34:187–199

Simon K, Xu J, Kim C, Skrynnikov NR (2005) Estimating the accuracy of protein structures using residual dipolar couplings. J Biomol NMR 33:83–93

Skrynnikov NR, Goto NK, Yang DW, Choy WY, Tolman JR, Mueller GA, Kay LE (2000) Orienting domains in proteins using dipolar couplings measured by liquid-state NMR: differences in solution and crystal forms of maltodextrin binding protein loaded with beta-cyclodextrin. J Mol Biol 295:1265–1273

Slupsky CM, Kay CM, Reinach FC, Smillie LB, Sykes BD (1995) Calcium-induced dimerization of troponin C: mode of interaction and use of trifluoroethanol as a denaturant of quaternary structure. Biochemistry 34:7365–7375

Sprangers R, Kay LE (2007) Quantitative dynamics and binding studies of the 20S proteasome by NMR. Nature 445:618–622

Strey R, Schomacker R, Roux D, Nallet F, Olsson U (1990) Dilute lamellar and L_3 phases in the binary water-C12E5 system. J Chem Soc Faraday T 86:2253–2261

Stubenrauch C, von Klitzing R (2003) Disjoining pressure in thin liquid foam and emulsion films—new concepts and perspectives. J Phys Condens Matt 15:R1197–R1232

Suh JY, Cai ML, Clore GM (2008) Impact of phosphorylation on structure and thermodynamics of the interaction between the N-terminal domain of enzyme I and the histidine phosphocarrier protein of the bacterial phosphotransferase system. J Biol Chem 283:18980–18989

Syvitski RT, Burnell EE (2000) Dipole-induced ordering in nematic liquid crystals. II. The elusive holy grail. J Chem Phys 113:3452–3465

Takino T, Nakada M, Miyamori H, Yamashita J, Yamada KM, Sato H (2003) CrkI adapter protein modulates cell migration and invasion in glioblastoma. Cancer Res 63:2335–2337

Terzis AF, Photinos DJ (1994) Electrostatic interactions in liquid crystals: ordering of rigid solutes in nematic solvents. Mol Phys 83:847–865

Tjandra N, Bax A (1997) Direct measurement of distances and angles in biomolecules by NMR in a dilute liquid crystalline medium. Science 278:1111–1114

Tollinger M, Forman-Kay JD, Kay LE (2002) Measurement of side-chain carboxyl pK_a values of glutamate and aspartate residues in an unfolded protein by multinuclear NMR spectroscopy. J Am Chem Soc 124:5714–5717

Tolman JR, Al-Hashimi HM, Kay LE, Prestegard JH (2001) Structural and dynamic analysis of residual dipolar coupling data for proteins. J Am Chem Soc 123:1416–1424

Tzakos AG, Grace CRR, Lukavsky PJ, Riek R (2006) NMR techniques for very large proteins and RNAs in solution. Annu Rev Biophys Biomol Struct 35:319–342

Ulmer TS, Werner JM, Campbell ID (2002) SH3-SH2 domain orientation in Src kinases: NMR studies of Fyn. Structure 10:901–911

Ulmer TS, Ramirez BE, Delaglio F, Bax A (2003) Evaluation of backbone proton positions and dynamics in a small protein by liquid crystal NMR spectroscopy. J Am Chem Soc 125:9179–9191

Valafar H, Prestegard JH (2004) REDCAT: a residual dipolar coupling analysis tool. J Magn Reson 167:228–241

van Dam L, Karlsson G, Edwards K (2006) Morphology of magnetically aligning DMPC/DHPC aggregates-perforated sheets, not disks. Langmuir 22:3280–3285

van Rossum BJ, Castellani F, Pauli J, Rehbein K, Hollander J, de Groot HJM, Oschkinat H (2003) Assignment of amide proton signals by combined evaluation of HN, NN and HNCA MAS-NMR correlation spectra. J Biomol NMR 25:217–223

Vega MC, Martínez JC, Serrano L (2000) Thermodynamic and structural characterization of Asn and Ala residues in the disallowed II' region of the Ramachandran plot. Protein Sci 9:2322–2328

von Berlepsch H, de Vries R (2000) Weakly charged lamellar bilayer system: interplay between thermal undulations and electrostatic repulsion. Eur Phys J E1:141–152

Walsh JD, Meier K, Ishima R, Gronenborn AM (2010) NMR studies on domain diffusion and alignment in modular GB1 repeats. Biophys J 99:2636–2646

Wang W, Weng J, Zhang X, Liu M, Zhang M (2009) Creating conformational entropy by increasing interdomain mobility in ligand binding regulation: a revisit to N-terminal tandem PDZ domains of PSD-95. J Am Chem Soc 131:787–796

Wardle KE, Carlson E, Henderson D, Rowley RL (2004) Molecular-dynamics simulation of the effect of ions on a liquid–liquid interface for a partially miscible mixture. J Chem Phys 120:7681–7688

Watanabe T, Tsuda M, Makino Y, Ichihara S, Sawa H, Minami A, Mochizuki N, Nagashima K, Tanaka S (2006) Adaptor molecule Crk is required for sustained phosphorylation of Grb2-associated binder 1 and hepatocyte growth factor-induced cell motility of human synovial sarcoma cell lines. Mol Cancer Res 4:499–510

Wong V, Case DA, Szabo A (2009) Influence of the coupling of interdomain and overall motions on NMR relaxation. Proc Natl Acad Sci USA 106:11016–11021

Wu B, Petersen M, Girard F, Tessari M, Wijmenga SS (2006) Prediction of molecular alignment of nucleic acids in aligned media. J Biomol NMR 35:103–115

Yang DW, Nagayama K (1996) A sensitivity-enhanced method for measuring heteronuclear long-range coupling constants from the displacement of signals in two 1D subspectra. J Magn Reson Ser A 118:117–121

Zangi R, Engberts JBFN (2005) Physisorption of hydroxide ions from aqueous solution to a hydrophobic surface. J Am Chem Soc 127:2272–2276

Zhang YB, Zuiderweg ERP (2004) The 70-kDa heat shock protein chaperone nucleotide-binding domain in solution unveiled as a molecular machine that can reorient its functional subdomains. Proc Natl Acad Sci USA 101:10272–10277

Zhang YJ, Oh H, Burton RA, Burgner JW, Geahlen RL, Post CB (2008) Tyr130 phosphorylation triggers Syk release from antigen receptor by long-distance conformational uncoupling. Proc Natl Acad Sci USA 105:11760–11765

Zheng J, Li LY, Chen SF, Jiang SY (2004) Molecular simulation study of water interactions with oligo (ethylene glycol)-terminated alkanethiol self-assembled monolayers. Langmuir 20:8931–8938

Zheng J, Li LY, Tsao HK, Sheng YJ, Chen SF, Jiang SY (2005) Strong repulsive forces between protein and oligo (ethylene glycol) self-assembled monolayers: A molecular simulation study. Biophys J 89:158–166

Zweckstetter M (2006) Prediction of charge-induced molecular alignment: residual dipolar couplings at pH 3 and alignment in surfactant liquid crystalline phases. Eur Biophys J Biophys Lett 35:170–180

Zweckstetter M (2008) NMR: prediction of molecular alignment from structure using the PALES software. Nat Protoc 3:679–690

Zweckstetter M, Bax A (2000) Prediction of sterically induced alignment in a dilute liquid crystalline phase: Aid to protein structure determination by NMR. J Am Chem Soc 122: 3791–3792

Zweckstetter M, Hummer G, Bax A (2004) Prediction of charge-induced molecular alignment of biomolecules dissolved in dilute liquid-crystalline phases. Biophys J 86:3444–3460