Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structural characterization of gephyrin by AFM and SAXS reveals a mixture of compact and extended states

Bodo Sander,[a] Giancarlo Tria,[b,c] Alexander V. Shkumatov,[b] Eun-Young Kim,[d] J. Günter Grossmann,[e] Ingrid Tessmer,[a] Dmitri I. Svergun[b] and Hermann Schindelin[a]*

[a]Structural Biology, Rudolf Virchow Center for Experimental Biomedicine, University of Würzburg, Josef-Schneider-Strasse 2, 97080 Würzburg, Germany, [b]Hamburg Outstation, European Molecular Laboratory c/o DESY, Notkestrasse 85, 22603 Hamburg, Germany, [c]Centre for Bioinformatics, University of Hamburg, Bundesstrasse 43, 20146 Hamburg, Germany, [d]Department of Biochemistry and Cell Biology, Stony Brook University, New York, NY 11794-5215, USA, and [e]Institute of Integrative Biology, Structural and Chemical Biology, University of Liverpool, Crown Street, Liverpool L69 7ZB, England

Correspondence e-mail: hermann.schindelin@virchow.uni-wuerzburg.de

Received 19 February 2013
Accepted 5 July 2013

Gephyrin is a trimeric protein involved in the final steps of molybdenum-cofactor (Moco) biosynthesis and in the clustering of inhibitory glycine and $GABA_A$ receptors at postsynaptic specializations. Each protomer consists of stably folded domains (referred to as the G and E domains) located at either terminus and connected by a proteolytically sensitive linker of ~150 residues. Both terminal domains can oligomerize in their isolated forms; however, in the context of the full-length protein only the G-domain trimer is permanently present, whereas E-domain dimerization is prevented. Atomic force microscopy (AFM) and small-angle X-ray scattering (SAXS) reveal a high degree of flexibility in the structure of gephyrin. The results imply an equilibrium between compact and extended conformational states in solution, with a preference for compact states. CD spectroscopy suggests that a partial compaction is achieved by interactions of the linker with the G and E domains. Taken together, the data provide a rationale for the role of the linker in the overall structure and the conformational dynamics of gephyrin.

1. Introduction

Gephyrin is a multifunctional protein that is involved in molybdenum-cofactor (Moco) biosynthesis as well as in the organization of the postsynaptic network at inhibitory synapses (Fritschy *et al.*, 2008; Tretter *et al.*, 2012). Structurally, gephyrin consists of two terminal domains connected by a linker. In the context of synaptic inhibition, the C-terminally located E domain of gephyrin binds to the β-subunit of glycine receptors with high affinity (Kim *et al.*, 2006) as well as to several subunits of the $GABA_A$ receptors (Kowalczyk *et al.*, 2013; Maric *et al.*, 2011; Mukherjee *et al.*, 2011; Tretter *et al.*, 2008, 2011), although the affinity is moderate in the latter cases. These interactions increase receptor dwell times at synaptic specializations by reducing receptor lateral diffusion rates (Dumoulin *et al.*, 2009; Meier *et al.*, 2001; Mukherjee *et al.*, 2011). In addition, gephyrin also forms complexes with proteins that aid in its tethering to synapses (neuroligin 2/4 and collybistin) as well as with proteins involved in the organization of the cytoskeleton (microtubules, profilin and Mena/VASP) or in cellular transport processes (dynein light chains 1 and 2 and KIF5) (Papadopoulos & Soykan, 2011; Tretter *et al.*, 2012).

The crystal structures of the terminal G and E domains of gephyrin revealed that they adopt the same architecture and oligomeric states as their bacterial orthologues MogA and MoeA, which like gephyrin are involved in Moco biosynthesis (Kim *et al.*, 2006; Schwarz *et al.*, 2001; Sola *et al.*, 2001, 2004). Specifically, the G domain (GephG) trimerizes, whereas the E domain (GephE) dimerizes, suggesting that gephyrin is able to form a hexagonal scaffold (Fig. 1) *via* the simultaneous utilization of both oligomerization interfaces (Kneussel & Betz, 2000; Xiang *et al.*, 2001). Although a definite proof for this model is still missing, this hypothesis is supported by the observation that the expression of variants lacking either domain or possessing mutations that interfere with the respective oligomerization interfaces disturbs gephyrin clustering (Calamai *et al.*, 2009; Saiyed *et al.*, 2007).

The formation of this scaffold, however, appears to require a trigger since full-length gephyrin has been reported to predominantly form trimers (Schrader *et al.*, 2004; Sola *et al.*, 2004). The structural properties of the central linker in the context of the full-length protein remain enigmatic. Trials to crystallize full-length gephyrin as well as truncated versions missing either domain or parts of the linker have so far failed, presumably owing to the linker being susceptible to proteolytic degradation. To gain insight into the overall structure of holo-gephyrin, including the relative arrangement of its domains and the linker, we used AFM and SAXS to describe its structure at nanometre resolution.

2. Materials and methods

2.1. Expression and purification

Rat full-length gephyrin splice variant P2 (residues 1–750; see Fig. 1) was cloned into pET28b vector (Novagen) *via* the *Nhe*I/*Hin*dIII restriction sites. Mutant constructs were created either by introducing a stop codon after residue 181 in the case of the GephG construct (residues 1–181) or by overlap mutagenesis for Geph-ΔL (lacking residues 187–331). Expression and purification of GephE (residues 332–750), gephyrin, GephG and Geph-ΔL were performed as described previously (Kim *et al.*, 2006; Schrader *et al.*, 2004) with modifications. A shallow salt gradient during anion-exchange chromatography (20 m*M* HEPES, 1 m*M* EDTA, 5 m*M* β-mercaptoethanol, 200–360 m*M* NaCl over 35 MonoQ 10/100 GL column volumes) allowed improved separation; in particular, higher gephyrin oligomers could be separated from

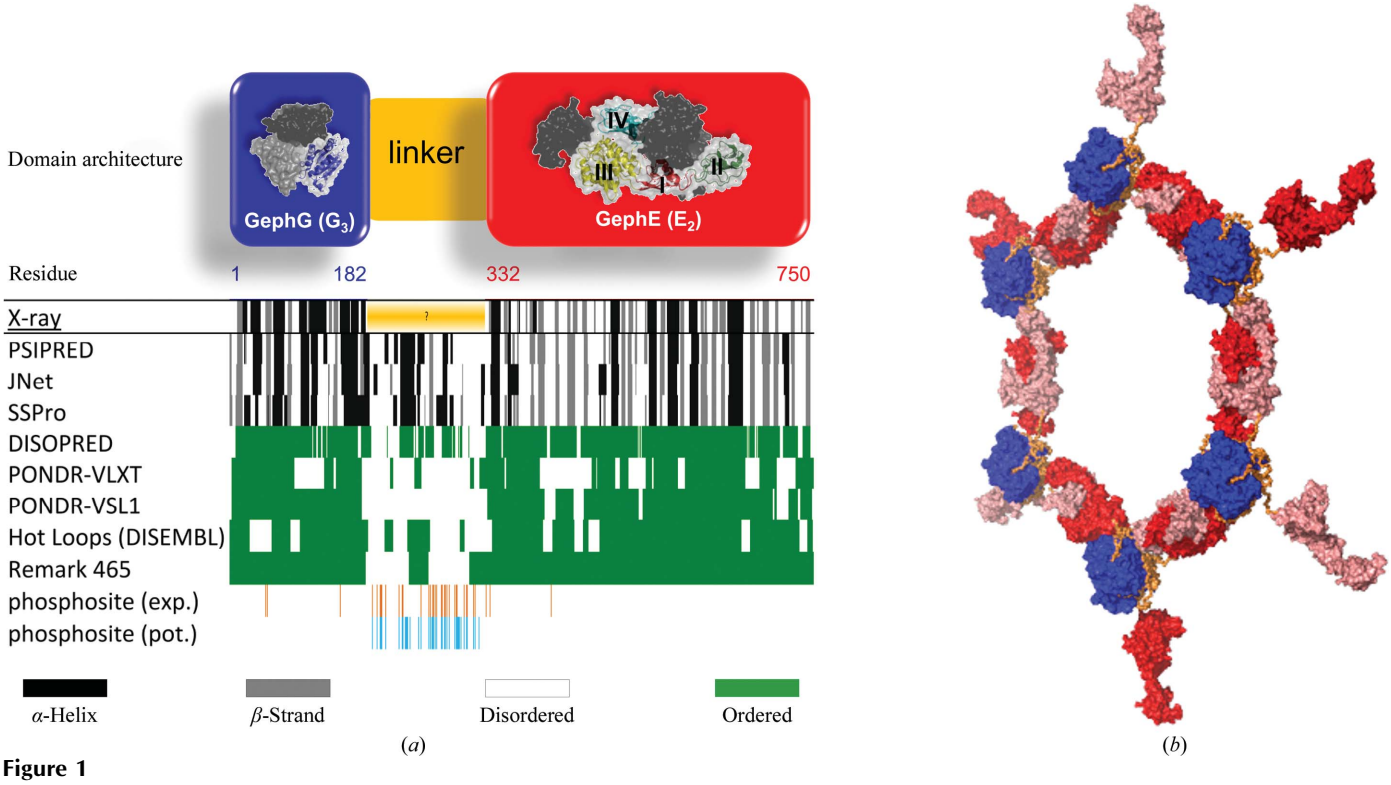


Figure 1
Domain architecture of gephyrin. (*a*) Predictions of secondary structure, disordered regions and phosphorylation sites. *PSIPRED*, *JNet* and *SSpro* accurately predict the secondary structure of the terminal domains GephG and GephE as derived from the crystal structures ('X-ray') with 75–78% accuracy. Secondary-structure elements are shown on top in line with residue numbers. GephG forms trimers and GephE dimers, with Roman numerals indicating the subdomain arrangement in one monomer. Large parts of the linker are predicted to be disordered; green patches correspond to ordered regions and colourless regions to unstructured regions. Cyan bars for the linker region represent all Ser, Thr and Tyr residues and thus all potential phosphorylation sites, whereas orange bars represent experimentally verified phosphorylation sites. Note that the major discrepancy between the potential and actual phosphorylation sites coincides with a predicted helical segment in the secondary-structure predictions. (*b*) Representation of the proposed hexagonal lattice relying on the simultaneous use of GephG trimers (G_3; blue) and GephE dimers (E_2; one protomer in red and the other in salmon). This assembly, which could theoretically be continued indefinitely, is thought to build a bridge between bound glycine or $GABA_A$ receptors and the cytoskeleton.

gephyrin trimers. The corresponding native gels (12.5 m*M* Tris, 96 m*M* glycine, 3.5% acrylamide:bis-acrylamide 80:1) allowed fractions containing the trimer to be selectively pooled so that an almost complete depletion of higher oligomers was achieved even before the size-exclusion chromatography step, after which a homogenous sample was obtained (see Supplementary Fig. S1[1] for details). All purified proteins were concentrated to 10–15 mg ml^{-1}, with the exception of GephG (5–10 mg ml^{-1}), as determined by UV absorption at 280 nm using calculated extinction coefficients of 30 660, 28 940, 21 680 and 7240 M^{-1} cm^{-1} for full-length gephyrin, Geph-ΔL, GephE and GephG, respectively. Proteins were flash-frozen for storage in liquid nitrogen, immediately prior to AFM or SAXS experiments, the samples were thawed and subjected to a Superdex 200 10/300 GL column (equilibrated in 20 m*M* HEPES pH 8, 250 m*M* NaCl, 1 m*M* EDTA, 5 m*M* β-mercaptoethanol, referred to as 'gephyrin buffer').

2.2. AFM

Protein samples were diluted in gephyrin buffer to a final concentration of 5–10 n*M* and immediately deposited onto freshly cleaved mica, rinsed with deionized water and dried in a gentle stream of nitrogen. Image collection was conducted on a Molecular Force Probe MFP-3D-BIO atomic force microscope (Asylum Research) in oscillating mode using Olympus OMCL-AC240 silicon probes with spring constants of ∼2 N m^{-1} and resonance frequencies of ∼70 kHz. 2 × 2 µm images were captured at a scan rate of 0.5 Hz and a resolution of 1024 × 1024 pixels. AFM experiments on wild-type and mutant gephyrin (Geph-ΔL; see §2.1) were carried out in triplicate.

AFM images were flattened to third order using the *Igor Pro*-based *MFP* software (Asylum Research). Peak volumes were measured using *ImageSXM* (S. Barret, University of Liverpool) at the level of individual molecules. The unstructured linker region of gephyrin was often not resolved in the topographical AFM images owing to its low height. Wherever connecting linker structures were not resolved, volumes of individual domains were summed up to obtain the total particle AFM volume (see, for example, the pool of AFM structures in Supplementary Fig. S2). Protein molecular masses were derived from the AFM volumes by comparison with a standard linear relationship obtained using calibration proteins (Roth *et al.*, 2012). The molecular mass was derived from the centre positions of Gaussian fits to the distributions of measured volumes (see Supplementary Fig. S2) using the *Origin* software. Error ranges were determined as the standard deviation given by the width of the Gaussians. Typically, molecular masses derived from AFM volume analysis are accurate to within 10% (Ratcliff & Erie, 2001). All particles with a molecular mass consistent with a gephyrin trimer (or within one standard deviation, SD, from the Gaussian centre) were chosen for analysis of the maximum diameter ($D_{\max}$).

Particle dimensions ($D_{\max}$ and peak heights, $h_{\rm gephyrin}$) were measured using *ImageSXM* and corrected for AFM tip-induced topography convolutions as previously described (Winzer *et al.*, 2012). Briefly, the widths of DNA fragments were measured with the same AFM probe as used for imaging of the gephyrin samples before and after gephyrin analysis. DNA sections were described by assuming a box shape with a width of 20 Å (the theoretical width of the DNA double helix) and height $h_{\rm DNA}$ as measured from the images using the section tool in the *MFP* software. DNA widths were measured using *ImageJ*. Comparison of the theoretical DNA width ($D_{\rm DNA}$ = 20 Å) and the measured width ($W_{\rm DNA}$) provides the radius of curvature of the employed AFM tip ($r_{\rm tip}$),

$$r_{\rm tip} \simeq \frac{(W_{\rm DNA} - D_{\rm DNA})^2}{8h_{\rm DNA}}. \quad (1)$$

$r_{\rm tip}$ was calculated by intrapolation of the DNA images for each gephyrin deposition, resulting in $r_{\rm tip}$ values of 6.6, 7.7 and 11.1 nm for the wild-type protein and of 5.1, 5.3 and 5.5 nm for Geph-ΔL. Finally, the contribution of the AFM tip to particle dimensions in the AFM images can be calculated and subtracted using (2) to obtain estimates of the true $D_{\max}$ values ($D_{\rm max,corrected}$),

$$D_{\rm max,corrected} = D_{\rm max,measured} - 2(2h_{\rm gephyrin}r_{\rm tip} - h_{\rm gephyrin}^2)^{1/2}. \quad (2)$$

All $D_{\max}$ values reported here have been corrected for AFM tip effects. For analysis, particles in the images were grouped according to the number of segments visible, from one segment for globular particles to four segments for highly extended molecular assemblies. Distributions of $D_{\max}$ data for the separate globular and extended species were Gaussian with $R^2 > 0.94$.

2.3. SAXS

Synchrotron X-ray scattering data for gephyrin were collected on the EMBL X33 beamline, DESY, Hamburg (Blanchet *et al.*, 2012; Roessle *et al.*, 2007) using a robotic sample changer (Round *et al.*, 2008). Initially, the data were reduced and processed using an automatic pipeline of scripts developed at EMBL Hamburg (Franke *et al.*, 2012). Gephyrin was prepared in gephyrin buffer (see §2.1) and measured at concentrations of 12, 7.5, 3 and 1.5 mg ml^{-1}. SAXS data were recorded at 283 K using a PILATUS 1M detector (DECTRIS, Baden, Switzerland) at a sample-to-detector distance of 2.7 m and a wavelength of 1.5 Å. This setup covers a momentum-transfer range of $0.01 < s < 0.6$ Å^{-1} [$s = 4\pi\sin(\theta)/\lambda$, where 2θ is the scattering angle]. Data recorded at 3 and 12 mg ml^{-1} were merged with *PRIMUS* (Konarev *et al.*, 2003; Petoukhov *et al.*, 2012). The forward scattering $I(0)$ and the radius of gyration ($R_{\rm g}$) were calculated using the Guinier approximation, assuming that at very small angles ($s < 1.3/R_{\rm g}$) the intensity is represented as $I(s) = I(0)\exp[-(sR_{\rm g})^2/3]$. The pair-distance distribution function $P(r)$, from which the maximum particle dimension ($D_{\max}$) and $R_{\rm g}$ were estimated, was computed using *GNOM* (Svergun, 1992). The molecular mass was derived

[1] Supplementary material has been deposited in the IUCr electronic archive (Reference: DZ5279). Services for accessing this material are described at the back of the journal.

from (i) the excluded volume of the hydrated particle using the Porod invariant (Petoukhov *et al.*, 2012) and (ii) the excluded volumes of *DAMMIN* models (Petoukhov *et al.*, 2012; Svergun, 1999) without symmetry.

Ab initio models using low-resolution data in the range $0.012 < s < 0.17$ Å^{-1} were created with *DAMMIN* (Svergun, 1999) and its faster version *DAMMIF* (Franke & Svergun, 2009). Both algorithms generate bead models yielding a scattering profile with the lowest possible discrepancy (χ) to the experimental data while keeping the beads interconnected and the model compact. Alternative *ab initio* models were created with *GASBOR* (Svergun *et al.*, 2001) using an extended data range ($0.012 < s < 0.5$ Å^{-1}). *GASBOR* requires the number of amino acids in the asymmetric unit and each bead represents a residue. Ten independent *ab initio* reconstructions were performed and were then averaged using *DAMAVER* (Volkov & Svergun, 2003), which also provides a value of normalized spatial discrepancy (NSD) representing a measure of similarity among different models. The figures and Table 2 refer to those models with the lowest average NSD value.

Combined *ab initio* and rigid-body modelling was performed with *BUNCH* (Petoukhov & Svergun, 2005) by imposing threefold symmetry and using the available high-resolution structures of the individual GephG (PDB entry 1jlj; Schwarz *et al.*, 2001) and GephE (PDB entry 2fts; Kim *et al.*, 2006) domains. The program *BUNCH* models the missing peptide segments as a chain of dummy residues that are separated by 3.8 Å to mimic a C$^{\alpha}$ chain. However, *BUNCH* can only be used for single-chain modelling and thus is restricted to symmetric assemblies. Rigid-body models without symmetry constraints (scenarios I and II; see Supplementary Fig. S4 for further description) were created with *CORAL* (Petoukhov *et al.*, 2012). As *CORAL* imposes a limit for the maximum linker length (99 residues), an alanine residue at position 263 (the numbering includes the His tag of the pET28 vector) was defined as a dummy rigid body, thus allowing the linker of 150 residues to be modelled as two linker segments of 59 and 90 residues, respectively. Ten models were generated for each scenario, this time using data in the range $0.012 < s < 0.5$ Å^{-1}. After modelling, the theoretical scattering profiles were computed with *CRYSOL* (Svergun *et al.*, 1995), fitting the experimental data within the interval $0.012 < s < 0.5$ Å^{-1}, using constant subtraction to account for a possible sample/buffer mismatch (Petoukhov *et al.*, 2012). *CRYSOL* fitting against the original data yielded the correct χ values. As deviations, especially in the small-angle region, were significant, the *AUTORG*-derived R_g (Petoukhov *et al.*, 2007) was imposed by a subsequent error reduction for s up to 0.020 Å^{-1}. However, even this approach did not yield satisfactory fits. Rigid-body models without imposed R_g and with the lowest χ value were taken as reference models and are presented in the figures.

Flexibility was assessed with the ensemble-optimization method (EOM; Bernadó *et al.*, 2007; Tria *et al.*, 2013), which assumes the coexistence of a range of conformations in solution for which an average scattering intensity fits the experimental SAXS data. In the first step, an enhanced version of *RANCH* (Petoukhov *et al.*, 2012) was used to create a pool of 950 000 independent models exhibiting either overall threefold or no symmetry for the linker and E domain. In this procedure residues belonging to the linker or the His tag were modelled to adopt 'native-like dihedral angles' instead of 'random-coil dihedral angles' (defined by default in *RANCH*), since ensembles for the former scenario yielded better fits in the small-angle region ($s < 0.03$ Å^{-1}). The theoretical scattering curve was then automatically computed for each model in the pool by *CRYSOL* (Svergun *et al.*, 1995). Afterwards, a genetic algorithm (*GAJOE*; Bernadó *et al.*, 2007; Petoukhov *et al.*, 2012) selected ensembles with varying numbers of conformers (from two to 40) by calculating the average theoretical profile and fitting it to the experimental SAXS data. *GAJOE* was repeated 100 times and the ensemble with the lowest discrepancy was reported as the best solution out of 100 final ensembles. Volume fractions of the corresponding models were further re-confirmed by *OLIGOMER* (Konarev *et al.*, 2006). In order to distinguish between EOM models that show extended and compact conformations, an R_g histogram was calculated using all models belonging to the selected ensembles. Models with R_g values above the average ($R_{g,average,pool} = 79.3$ Å) were classified as extended and models with values below the average as compact. SAXS methods and results are summarized in Tables 1 and 2.

Table 1
SAXS data analysis.

Data-collection parameters	
Instrument	EMBL X33 beamline, DORIS III storage ring (DESY, Hamburg), PILATUS 1M detector
Beam size at the detector (mm)	2 × 0.6
Wavelength (Å)	1.5
q-range (Å^{-1})	0.008–0.600
Exposure time (s)	8 × 15
Concentration range (mg ml^{-1})	1.5–12
Temperature (K)	283
Structural parameters	
R_g [real-space R_g from $P(r)$] (Å)	66.7 ± 2
R_g (from Guinier) (Å)	62.5 ± 2
D_{max} (Å)	240 ± 25
Porod volume estimate (Å^3)	369000 ± 5000
Dry volume calculated from sequence (Å^3)	305700
Molecular-mass determination	
Partial specific volume (cm^3 g^{-1})	0.739
Molecular mass M_r (Da)	
From Porod volume (V_p/1.6)	231000 ± 45000
From excluded volume (V_{ex}/2)	232000 ± 45000
From SAXS MoW	226000 ± 23000
Calculated for trimer from sequence	250536
Software employed	
Primary data reduction	*PRIMUS*
Data processing	*PRIMUS*
Ab initio analysis	*DAMMIN*/*DAMMIF*/*GASBOR*
Validation and averaging	*DAMAVER*
Rigid-body modelling	*BUNCH*/*CORAL*
Flexibility	*EOM*
Computation of model intensities	*CRYSOL*
Three-dimensional graphics representations	*PyMOL*

Table 2
SAXS model fits to experimental data.

NSD, normalized spatial discrepancy; asym. I and asym. II, rigid-body modelling of asymmetric scenarios I and II, respectively.

		Asymmetric			Symmetric†		
	Program	χ	NSD	R_g (Å)	χ	NSD	R_g (Å)
Ab initio	*DAMMIN*	0.81	0.72 ± 0.03	66.8	0.97	1.30 ±0.15	65.3
	DAMMIF	0.85	0.76 ± 0.03	66.9	1.55	1.37 ± 0.22	66.9
	GASBOR	1.36	1.98 ± 0.06	67.2	1.80	1.63 ± 0.27	64.6
Rigid body	*CORAL* (asym. I)	1.46	2.47 ± 0.03	80.6			
	CORAL (asym. II)	2.27	2.32 ± 0.08	79.8			
	BUNCH (symmetric)				1.50	1.52 ± 0.17	66.9
EOM	*EOM*‡			80.8	0.80		

† Symmetry refers to overall threefold symmetry. ‡ For EOM analysis a mixed pool of both symmetric and asymmetric linkers was created. Therefore, it is classified as neither purely symmetric nor purely asymmetric. It should be noted, that the Guinier approximation for the EOM ensemble curve using the same resolution range as for the experimental data (compare with inset in Fig. 4*b*) yields an R_g value of 65 Å.

2.4. Dynamic light scattering

Dynamic light-scattering (DLS) experiments of full-length gephyrin were carried out at a protein concentration of 1.5 mg ml^{-1} with a DynaPro Titan device (Wyatt) at 283 K in gephyrin buffer (see §2.1). The hydrodynamic radius (R_h) was determined with the *DYNAMICS* 6.7.3 software. The polydispersity was below 20% (data not shown).

2.5. Circular-dichroism spectroscopy

Circular-dichroism (CD) spectroscopy was conducted with a Jasco J-810 spectropolarimeter. Far-UV spectra from 190 to 260 nm were recorded at a scanning speed of 20 nm min^{-1} with a response time of 1 s and a bandwidth of 2 nm. The buffer was exchanged to 50 m*M* potassium phosphate with ultrafiltration units (Sartorius Vivaspin 500, Göttingen). The decrease in circular dichroism at 200 nm (bandwidth = 2 nm) was measured repeatedly (n = 5) as a function of temperature ($293 < T < 368$ K) to obtain the melting curves at a heating rate of 1 K min^{-1}.

3. Results

3.1. Bioinformatic analysis of gephyrin

Gephyrin consists of two terminally folded domains connected by a linker for which no structural information is available. As it is well known that full-length gephyrin is prone to degradation (Herweg & Schwarz, 2012), we compared secondary-structure and disorder predictions (Fig. 1). Six independent algorithms (Cole *et al.*, 2008; Kelley & Sternberg, 2009; Li *et al.*, 1999; Linding *et al.*, 2003; McGuffin *et al.*, 2000; Pollastri *et al.*, 2002; Ward *et al.*, 2004) consistently predict large parts of the linker to be unstructured (ranging from 56 to 100%). These results can partly be explained by the fact that the residues in the linker are biased towards a reduced content of hydrophobic, aromatic, bulky and order-promoting side chains and an increased content of residues with a higher solvation potential as well as disorder-promoting residues (compared with GephG and GephE; Supplementary Table S1; Vacic *et al.*, 2007). Notably, the predicted degree of order/disorder is not evenly distributed in the linker region and the following regions can be distinguished. The terminal segments of the linker of gephyrin (residues 181–222 and residues 281–332) are predicted to contain hardly any secondary structure. Half of the region between residues 242 and 280 is predicted to be structured, and the only consistently predicted segment of secondary structure (residues 224–241) forms an α-helix.

These results suggest a predominantly unstructured linker leading to a high degree of molecular plasticity as well as to the ability to adopt a rather extended overall conformation. To clarify the impact of potential flexibility on the relative arrangement of globular protein domains, we further characterized gephyrin using the complementary methods of AFM and SAXS.

3.2. AFM analysis

Single-molecule imaging by AFM offers a potent means for the direct analysis of possible conformational heterogeneity in the gephyrin assemblies (Lemaire *et al.*, 2006; Sacho *et al.*, 2008). Our AFM images showed that the samples were homogeneous in terms of oligomeric state: the AFM volumes displayed a Gaussian distribution with a centre value corresponding to a molecular mass of 219 ± 39 kDa (Supplementary Fig. S2) consistent with the theoretical mass of the gephyrin trimer of 250.5 kDa. Considering that linker segments are not always resolved in the images owing to their low height as well as their missing organization and hence do not always contribute to the total mass of the particles, one would indeed expect the mass to be slightly underestimated. However, the molecular mass of the sample particles is still significantly higher than that of a potential dimer (expected dimer mass of 167 kDa), confirming that the majority of gephyrin particles were trimeric, while only 1.9 ± 0.1% of the particles were consistent with a hexameric state.

With respect to their shape, however, the particles displayed highly diverse domain arrangements varying between compact and very extended assemblies (Fig. 2*a* and Supplementary Fig. S2). This indicates a high level of plasticity in the molecules, which is also reflected in the broad AFM-derived D_{max} distribution ranging from ∼200 to ∼500 Å (Fig. 2*b*). Particles were classified according to the number of segments that they displayed (for examples, see Fig. 2*a* and Supplementary Fig. S2), with the number of segments ranging from one to four: globular particles consisting of only one segment made up 15.6 ± 4.0% of the sample, while 29.5 ± 1.3% contained two segments, 37.2 ± 0.6% contained three segments and 17.8 ± 3.0% contained four segments (Fig. 2*a* and Supplementary Fig. S2). The large majority of the multi-cluster particles (two

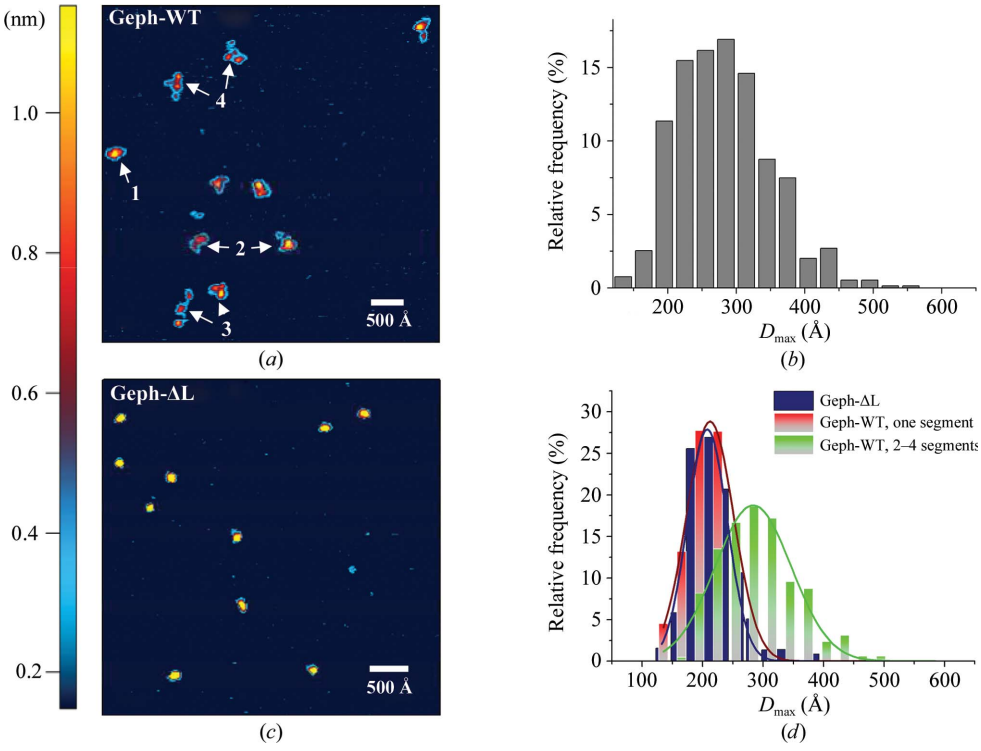


Figure 2
AFM analysis of gephyrin and a gephyrin variant with the linker missing (Geph-ΔL). (*a*) AFM micrograph displaying distinct classes of particles containing one to four segments as indicated in the figure (1–4). The height scale is indicated by the colour bar on the left-hand side and applies to (*a*) and (*c*). (*b*) Size distribution of the trimeric fraction of wild-type gephyrin (n = 413). (*c*) In contrast to gephyrin, Geph-ΔL displays a high level of homogeneity. (*d*) Separate size distributions of the compact and extended gephyrin populations (one segment and 2–4 segments, respectively) can be described by Gaussian curves centred at 212 and 283 Å, respectively. The Geph-ΔL mutant shows a size distribution that strongly resembles that of compact full-length gephyrin (centre of the Gaussian fit at 229 Å).

and three clusters) displayed a profile that could principally be superimposed with asymmetric as well as symmetric models. Among the population with four clusters, >90% are consistent with a quasi-symmetric assembly. Globular and extended conformers containing 2–4 clusters were analysed separately for particle size (Fig. 2*d*), resulting in D_{max} values of 212 and 283 Å, respectively.

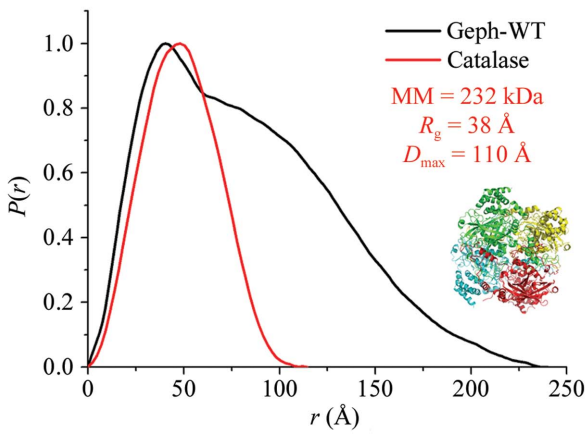
These results are in line with previous limited proteolysis studies, which also suggest that it is the linker that mediates heterogeneity. To test this hypothesis, a gephyrin construct missing the linker (Geph-ΔL) was analysed. The volume-derived molecular mass of 223.3 ± 8.7 kDa is in line with Geph-ΔL forming a trimer (expected molecular mass of 200 kDa). We attribute the slightly larger volume-derived molecular mass for Geph-ΔL compared with full-length gephyrin to the fact that some parts of the linker are not resolved in the AFM analysis of full-length gephyrin and the holo-gephyrin mass has therefore been underestimated, as outlined in the previous paragraph. Since virtually all Geph-ΔL particles appeared to be globular in the AFM micrographs (Fig. 2*c*) and their D_{max} values were very similar to those observed for the globular fraction of full-length gephyrin, one can conclude that the linker mediates an equilibrium between compact and extended states.

3.3. SAXS analysis

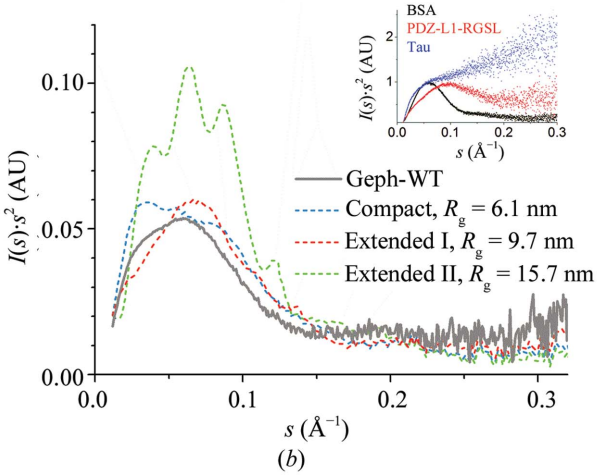
3.3.1. Overall characterization. The trimeric state of gephyrin (Schrader *et al.*, 2004) was confirmed by the SAXS-derived overall parameters. Molecular masses of 231 ± 45, 232 ± 45 and 226 ± 23 kDa were deduced from the Porod volume (Porod, 1982), excluded volume (Petoukhov *et al.*, 2012) and the SAXS MoW calculation (Fischer *et al.*, 2010), respectively (Table 1). These values are comparable to an expected trimer mass of 250.5 kDa. Initial clues about the shape of gephyrin can be derived from the interatomic pair-distance distribution, the $P(r)$ function, which suggests an anisometric shape for gephyrin: firstly, the $P(r)$ function is asymmetric as is typical for extended assemblies, and secondly, the parameters characterizing the overall dimensions, the maximum distance (D_{max}) and the radius of gyration (R_g), with values of ∼240 and ∼65 Å, respectively, are rather high for a molecule with the mass of gephyrin (Fig. 3*a*).

Consistently, *ab initio* models created with *DAMMIN*, *DAMMIF* and *GASBOR* without symmetry constraints yielded rather elongated shapes (Supplementary Fig. S3). However, it will be illustrated below that in the case of gephyrin shape reconstructions without constraints should be considered with great caution. The application of threefold symmetry, potentially justified by the oligomeric state of the protein, also led to unreliable and ambiguous results (Supplementary Fig. S3).

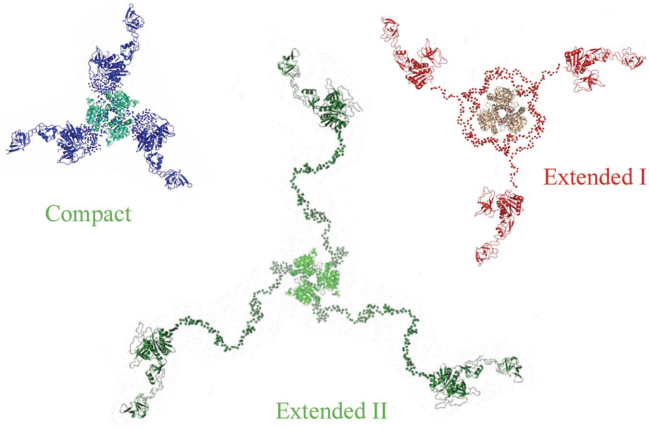
At the same time, several findings clearly indicate that the linker is not entirely flexible, potentially pointing to a considerable degree of compaction. Firstly, one might have expected an even larger D_{max} considering the 150-residue linker, which theoretically could expand the GephG–GephE distance by up to ∼450 Å for each protomer. Secondly, the R_g/R_h ratio of 0.93 (R_g as determined by SAXS and R_h by DLS) corresponds to a relatively compact assembly. For comparison, R_g/R_h ranges from ∼1.7 for anisometric polymers to ∼0.8 for a solid sphere. Finally, the bell-shaped Kratky plot of gephyrin suggests that the protein is predominantly folded with a rather low contribution from random coils (Fig. 3*b*). We compared the Kratky plot with those computed for three conformers of different compactness and found that the experimentally derived Kratky plot can roughly be considered as an intermediate between a very compact and a slightly extended state



(*a*)



(*b*)



(*c*)

Figure 3
SAXS analysis of gephyrin. (*a*) Normalized $P(r)$ functions of catalase (red) which forms a globular 232 kDa homotetramer (see ribbon diagram) and trimeric (250.5 kDa) gephyrin (black). The tail in the $P(r)$ function of gephyrin reveals its elongated shape. Compared with gephyrin (R_g = 67 Å and D_{max} = 250 Å), the corresponding values for catalase (R_g = 38 Å and D_{max} = 110 Å) are reduced by ∼50%. (*b*) Experimentally derived Kratky plot of gephyrin in comparison with simulated Kratky plots for three different extended conformers as shown in (*c*) and, in the inset, with globular BSA, natively unfolded tau (Shkumatov *et al.*, 2011) and PDZ-L1-RGSL, a two-domain protein with a flexible linker representing 30% of the total protein (Bielnicki *et al.*, 2011) in a normalized representation. (*c*) The conformers used for the Kratky plot comparison in (*b*).

(Figs. 3*b* and 3*c*). In fact, a peak broadening in Kratky plots similar to that of gephyrin was observed in simulated data of multidomain proteins connected by flexible linkers, which also displayed a tendency for a single peak in the Kratky plot that might point to a decoupling of the movements of GephG and GephE (Bernadó, 2010).

3.3.2. Combined *ab initio*/rigid-body modelling. Since high-resolution structures are available for ∼80% of the primary sequence, rigid-body modelling was performed with the GephG and GephE crystal structures, initially using *CORAL*, assuming no symmetry except for the trimeric GephG which mediates the trimerization of gephyrin. In accordance with the observed flexibility, however, rigid-body modelling turned out to be problematic: firstly, rigid-body models displayed a high normalized spatial discrepancy (NSD), reflecting the ambiguity of the solutions (Supplementary Fig. S4) and indicating that a single molecular conformation alone does not represent the experimental data adequately. Secondly, models could not be obtained which were simultaneously biologically relevant and yielded good fits over the complete data range. The best fits could be obtained with symmetric modelling using *BUNCH*; however, the obtained rigid triskelion does not explain the sample heterogeneity observed with AFM, indicating that rigid-body modelling assuming single rigid conformers is not the method of choice to describe the scattering pattern of gephyrin (Supplementary Fig. S4).

3.3.3. Ensemble-optimization method. The ensemble-optimization method (EOM; Bernadó *et al.*, 2007; Tria *et al.*, 2013) allows one to analyse the scattering data in terms of ensembles of conformers and thus appears to be the best choice for the analysis of gephyrin. In short, a large pool of conformers (half of them symmetric) was created. A genetic algorithm then selected sub-ensembles whose average theoretical scattering fits to the experimental SAXS data. Finally, the R_g histogram of the initial pool was compared with the corresponding histogram of 100 calculated ensembles (Fig. 4*a*), where models were classified as compact or extended according to their radii of gyration (see Fig. 4*a* and §2). This EOM analysis reveals two interesting characteristics of gephyrin: firstly, almost the entire R_g range of the initial pool (red line in Fig. 4*a*) is represented in the selected pool (red dashed line in Fig. 4*a*) indicative of a high degree of flexibility. Secondly, the R_g distribution for the selected ensemble displays a clear maximum at R_g = 58 Å, followed by a shoulder at R_g = 93 Å and a third rather broad and not very pronounced maximum at $R_g \simeq$ 140 Å. The lack of a more prominent peak for the extended states indicates a higher heterogeneity of this population.

The ensemble with the best fit (χ = 0.80; Fig. 4*b*) contains six conformers (Fig. 4*c*) and their R_g values mirror the peak positions of the R_g histogram. Four conformers of this ensemble are compact (R_g values of 54–63 Å) and contribute 63% to the scattering, while the remaining contribution of 37% stems from two extended conformers with R_g values of either 92 or 134 Å. A determination of the volume fractions with the program *OLIGOMER* is in line with the EOM results (χ = 0.80): the compact conformers contribute ∼65% to the

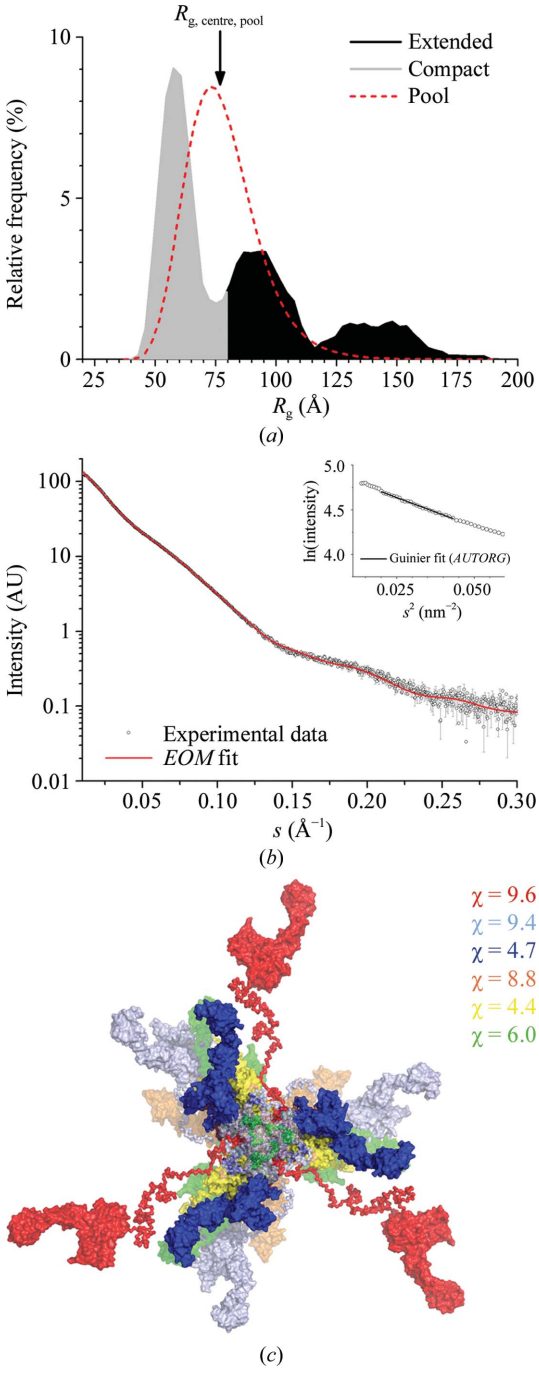


Figure 4
EOM analysis of gephyrin. (*a*) EOM R_g distribution of the initial random pool (red dashed line) and the selected ensemble (different grey shades) for full-length gephyrin. The broad distribution mirrors the conformational heterogeneity of the sample; however, there is a slight preference for compact states as indicated by the larger area under the curve for this fraction when compared with intermediate and extended conformers. (*b*) The simulated curve of the selected EOM ensemble is compared with the experimental curve with open circles and error bars. The Guinier plot in the inset reveals the structural integrity of the protein, with the line indicating the data points used for R_g and $I(0)$ determination with *GNOM*. (*c*) EOM ensemble of the gephyrin trimer in which different trimers are coloured differently. A higher degree of transparency for the yellow and the blue conformers indicates their smaller contributions to the total scattering. The individual conformers were also fitted against the experimental data and the obtained χ values are displayed in the corresponding colour.

total scattering and the extended structures contribute ∼35%. EOM analyses using exclusively asymmetric assemblies did not improve the fit to the data compared with that obtained for the mixed symmetric/asymmetric pool (results not shown).

3.4. CD spectroscopy

Compaction of a presumably unstructured linker can also result from interactions with folded domains (Babu *et al.*, 2012). To analyse the possible stabilization of either GephG or GephE by the linker, we recorded thermal unfolding curves of gephyrin *via* the decrease in the circular dichroism at 200 nm (Fig. 5). Melting curves were determined for the full-length protein as well as for the individual terminal domains. GephG (residues 1–181) in its isolated form displays a melting temperature T_m of 349.8 K, while isolated GephE unfolded earlier at a T_m of 332.0 K. In the context of full-length gephyrin, transitions at 353.9 and 335.5 K were observed, and these can be attributed to the unfolding of GephG and GephE, respectively. The measured melting temperatures reveal that both terminal domains are stabilized by 3–4 K in the presence of the linker.

4. Discussion

4.1. Oligomeric state

In this study, the structure of trimeric full-length gephyrin was analysed. While expression in *E. coli* results predominantly in a trimeric assembly, which accounts for more than 75% of the total protein, several higher order oligomers (presumably multiples of trimers) were also obtained. Higher oligomers could be separated from the trimeric fraction by anion-exchange chromatography utilizing a shallow salt gradient and native PAGE gels so that the higher oligomers did not affect data analysis (Supplementary Fig. S1). While the higher oligomers could be separated easily from the trimer, they could not be purified individually to the same homogeneity as the trimers (data not shown). In the context of this study, we focused on the prevalent trimeric form of full-length gephyrin, which presumably represents the building block for the higher oligomers (hexamers and larger) that were also observed when gephyrin was overexpressed in eukaryotes (Herweg & Schwarz, 2012; Saiyed *et al.*, 2007).

4.2. Impact of the mostly unstructured linker

The 150-residue predominantly unstructured linker between the two globular domains, GephG and GephE, renders the gephyrin trimer rather flexible, preventing structural analysis of full-length gephyrin by crystallography. In line with the observed proteolytic sensitivity of the linker (Herweg & Schwarz, 2012), bioinformatics analyses support the conjecture that large segments of the linker are intrinsically disordered. Consistent with this, 70% of all serines, threonines and tyrosines (a total of 40 residues) in the linker have been verified as phosphorylation targets (Herweg & Schwarz, 2012; Huttlin *et al.*, 2010), which are usually solvent-exposed to

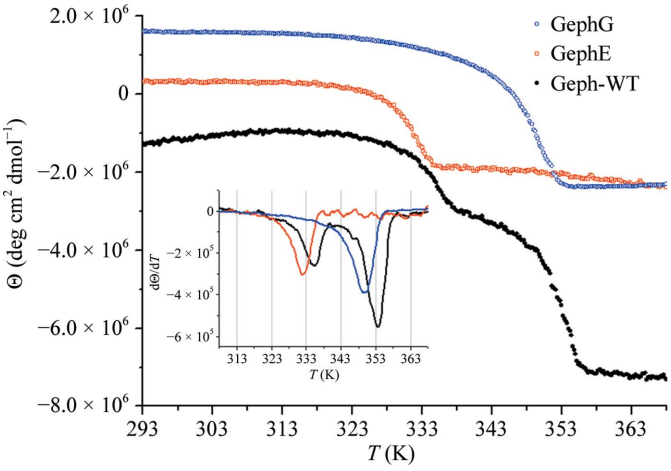


Figure 5
Thermal unfolding of gephyrin (black) and its isolated terminal domains (G domain in blue and E domain in red) monitored by changes in the circular dichroism. The minima of the resulting derivatives (inset) reveal that the melting temperatures of GephG and GephE are increased in the context of the full-length protein.

provide easily accessible binding epitopes for the respective kinases (Dunker *et al.*, 2002).

Single-molecule analysis of trimeric gephyrin with AFM clearly demonstrates a significant conformational variety, yielding assemblies ranging from globular particles with a single segment (species 1 in Fig. 2*a*) to those with different degrees of extended shape (species 2–4 in Fig. 2*a*). This structural flexibility translates into a broad $D_{\max}$ distribution similar to that observed in the SAXS ensemble-optimization (EOM) analysis. A gephyrin variant lacking virtually the complete linker showed almost identical particle dimensions ($D_{\max}$) to the globular single-cluster conformers of wild-type gephyrin (Figs. 2*c* and 2*d*), thus demonstrating that it is the linker that causes the multitude of conformational landscapes observed by SAXS and AFM. In summary, the AFM-derived $D_{\max}$ distribution and the particle classification, which are in agreement with the SAXS results, suggest an equilibrium between compact and extended conformations.

As expected from bioinformatics and our AFM data, conventional SAXS data analysis, either by shape reconstruction or rigid-body modelling assuming a single conformation, cannot accurately model this multi-domain assembly (Table 2 and Supplementary Fig. S4). Hints to the structural heterogeneity of gephyrin in SAXS also came from the Kratky plot (single peak in Fig. 3*b*) and missing interdomain correlation peaks in the $P(r)$ function as well as the relatively high χ values and the weak reproducibility in the case of individual rigid-body models (as indicated by NSD values significantly higher than 1; Table 2 and Supplementary Fig. S4; Bernadó, 2010).

Ab initio reconstructions assuming single conformers and no symmetry constraints resulted in elongated models. AFM and SAXS analyses further indicated that gephyrin is a rather flat molecule and thus represents a special case in which *ab initio* reconstructions without symmetry restrictions fail (see test case 12 in Fig. 4 of Volkov & Svergun, 2003). The present study and results are therefore interesting since cases of extremely flat molecules for which *ab initio* analysis without symmetry constraints fails have not yet been described in practice, except for the analysis of immunoglobulin M (Volkov *et al.*, 2003). Structural flexibility was further assessed with the advanced v.2.0 of *EOM* (Tria *et al.*, 2013) and the resulting models are consistent with an elastic linker that allows trimeric gephyrin to switch between compact and very elongated conformations. More importantly, the EOM results indicate a compaction of a significant fraction of the gephyrin molecules.

This partial compaction can be explained by the fact that some parts of the linker are not flexible and seem to interact with both GephG and GephE. This conclusion is supported by the increased thermal stabilization of both terminal domains observed here with CD spectroscopy and recently by differential scanning calorimetry (Herweg & Schwarz, 2012). Interactions with the linker and the terminal globular domains might explain at least in part why the radius of gyration is considerably smaller than would be expected for a gephyrin trimer with fully extended linkers, where R_g values larger than 100 Å would be accessible. In case of GephE the interaction with the linker not only results in thermal stabilization but was also found to be responsible for blocking E-domain-mediated dimerization (Bedet *et al.*, 2006), thus explaining why GephE without the linker dimerizes and full-length gephyrin predominantly forms trimers after expression in *E. coli*.

4.3. Implications for gephyrin structure–function relationships

The structural data presented here suggest that both primary functions of gephyrin, Moco biosynthesis and clustering of inhibitory neurotransmitter receptors, benefit from its conformational flexibility. (i) Gephyrin catalyses the final two steps in the pathway of Moco biosynthesis, namely the adenylation of molybdopterin (MPT) at the active site of GephG followed by the removal of adenosine monophosphate (AMP) and copper, which is replaced by molybdenum at the active site of GephE (Schwarz *et al.*, 2009). In a recent study, it could be demonstrated that the linker region of gephyrin modulates its enzymatic activity and that close proximity of the active sites of GephG and GephE on its own is not sufficient for proper Moco biosynthesis (Belaidi & Schwarz, 2013). (ii) With respect to receptor clustering, it has been proposed that oligomerization *via* GephE in addition to that by GephG paves the way for the formation of a hexagonal lattice just beneath the postsynaptic membrane of inhibitory synapses, allowing the tethering of glycine and $GABA_A$ receptors on one side and binding to cytoskeletal elements (Xiang *et al.*, 2001) on the other. Detailed regulatory mechanisms to overcome the blockage of GephE dimerization have not yet been established, but a key role for the comparatively long linker segment of gephyrin (consisting of 136–196 residues depending on the splice variant) is plausible in the light of an emerging body of literature describing intrinsically disordered proteins. Protein segments can serve as important and versatile molecular-control devices that are able to sustain an

abundance of conformations and transitions which are important for function, including molecular and cellular regulation (Babu *et al.*, 2012). In fact, the tightly controlled influence of conformational flexibility of disordered regions allows the exposure of short linear motifs, which enable specific protein–protein interactions (Davey *et al.*, 2012) or post-translational modifications.

In the case of gephyrin, interactions with dynein light chains 1 and 2, ena/VASP and collybistin involve segments of the linker region (Bausen *et al.*, 2006; Maas *et al.*, 2006; Tyagarajan, Ghosh, Harvey *et al.*, 2011), and collybistin, for example, has been regarded as a binding partner crucial for the clustering of gephyrin. In addition, gephyrin phosphorylation has attracted a great deal of attention, resulting in studies on the impact of phosphorylation on the clustering properties of gephyrin (Bausen *et al.*, 2010; Kuhse *et al.*, 2012; Tyagarajan, Ghosh, Yevenes *et al.*, 2011; Zita *et al.*, 2007). For example, it has been demonstrated that mutations in the second half of the linker influence clustering of gephyrin (Kuhse *et al.*, 2012; Tyagarajan, Ghosh, Yevenes *et al.*, 2011). Hence, it will be necessary to identify auto-inhibitory linker segments that interact with the terminal domains and lead to compaction of the resulting gephyrin assemblies.

This work was supported by the Deutsche Forschungsgemeinschaft (Rudolf Virchow Center for Experimental Biomedicine, FZ 82 to HS and IT), Schi 425/8-1 and SFB487 C7 (to HS). BS was supported by the BIGSS (BioMedTec International Graduate School of Science) Lead Structures of Cell Function, part of the Elitenetzwerk Bavaria. The work of AVS and DIS was supported by Bundesministerium für Bildung und Forschung (grant BIOSCAT, contract No. 05K12YE1) and European Union FP7 e-Infrastructure (grant WeNMR, contract No. 261572). We thank the EMBL Hamburg for synchrotron beamtime allocation at DESY (Hamburg, Germany), Eun Young Lee for cloning of the gephyrin wild-type construct and Dr Bernhard Fröhlich as well as Christian Weinberger for computational assistance. STFC Daresbury Laboratory is also acknowledged where seminal SAXS experiments on gephyrin were carried out at SRS station 2.1.

References

Babu, M. M., Kriwacki, R. W. & Pappu, R. V. (2012). *Science*, **337**, 1460–1461.

Bausen, M., Fuhrmann, J. C., Betz, H. & O'Sullivan, G. A. (2006). *Mol. Cell. Neurosci.* **31**, 376–386.

Bausen, M., Weltzien, F., Betz, H. & O'Sullivan, G. A. (2010). *Mol. Cell. Neurosci.* **44**, 201–209.

Bedet, C., Bruusgaard, J. C., Vergo, S., Groth-Pedersen, L., Eimer, S., Triller, A. & Vannier, C. (2006). *J. Biol. Chem.* **281**, 30046–30056.

Belaidi, A. A. & Schwarz, G. (2013). *Biochem. J.* **450**, 149–157.

Bernadó, P. (2010). *Eur. Biophys. J.* **39**, 769–780.

Bernadó, P., Mylonas, E., Petoukhov, M. V., Blackledge, M. & Svergun, D. I. (2007). *J. Am. Chem. Soc.* **129**, 5656–5664.

Bielnicki, J. A., Shkumatov, A. V., Derewenda, U., Somlyo, A. V., Svergun, D. I. & Derewenda, Z. S. (2011). *J. Biol. Chem.* **286**, 35163–35175.

Blanchet, C. E., Zozulya, A. V., Kikhney, A. G., Franke, D., Konarev, P. V., Shang, W., Klaering, R., Robrahn, B., Hermes, C., Cipriani, F., Svergun, D. I. & Roessle, M. (2012). *J. Appl. Cryst.* **45**, 489–495.

Boze, H., Marlin, T., Durand, D., Pérez, J., Vernhet, A., Canon, F., Sarni-Manchado, P., Cheynier, V. & Cabane, B. (2010). *Biophys. J.* **99**, 656–665.

Calamai, M., Specht, C. G., Heller, J., Alcor, D., Machado, P., Vannier, C. & Triller, A. (2009). *J. Neurosci.* **29**, 7639–7648.

Cole, C., Barber, J. D. & Barton, G. J. (2008). *Nucleic Acids Res.* **36**, W197–W201.

Davey, N. E., Van Roey, K., Weatheritt, R. J., Toedt, G., Uyar, B., Altenberg, B., Budd, A., Diella, F., Dinkel, H. & Gibson, T. J. (2012). *Mol. Biosyst.* **8**, 268–281.

Dumoulin, A., Triller, A. & Kneussel, M. (2009). *Front Mol. Neurosci.* **2**, 28.

Dunker, A. K., Brown, C. J., Lawson, J. D., Iakoucheva, L. M. & Obradović, Z. (2002). *Biochemistry*, **41**, 6573–6582.

Fischer, H., de Oliveira Neto, M., Napolitano, H. B., Polikarpov, I. & Craievich, A. F. (2010). *J. Appl. Cryst.* **43**, 101–109.

Franke, D., Kikhney, A. G. & Svergun, D. I. (2012). *Nucl. Instrum. Methods Phys. Res. A*, **689**, 52–59.

Franke, D. & Svergun, D. I. (2009). *J. Appl. Cryst.* **42**, 342–346.

Fritschy, J. M., Harvey, R. J. & Schwarz, G. (2008). *Trends Neurosci.* **31**, 257–264.

Herweg, J. & Schwarz, G. (2012). *J. Biol. Chem.* **287**, 12645–12656.

Huttlin, E. L., Jedrychowski, M. P., Elias, J. E., Goswami, T., Rad, R., Beausoleil, S. A., Villén, J., Haas, W., Sowa, M. E. & Gygi, S. P. (2010). *Cell*, **143**, 1174–1189.

Kelley, L. A. & Sternberg, M. J. (2009). *Nature Protoc.* **4**, 363–371.

Kim, E. Y., Schrader, N., Smolinsky, B., Bedet, C., Vannier, C., Schwarz, G. & Schindelin, H. (2006). *EMBO J.* **25**, 1385–1395.

Kneussel, M. & Betz, H. (2000). *Trends Neurosci.* **23**, 429–435.

Konarev, P. V., Petoukhov, M. V., Volkov, V. V. & Svergun, D. I. (2006). *J. Appl. Cryst.* **39**, 277–286.

Konarev, P. V., Volkov, V. V., Sokolova, A. V., Koch, M. H. J. & Svergun, D. I. (2003). *J. Appl. Cryst.* **36**, 1277–1282.

Kowalczyk, S., Winkelmann, A., Smolinsky, B., Förstera, B., Neundorf, I., Schwarz, G. & Meier, J. C. (2013). *Eur. J. Neurosci.* **37**, 544–554.

Kuhse, J., Kalbouneh, H., Schlicksupp, A., Mükusch, S., Nawrotzki, R. & Kirsch, J. (2012). *J. Biol. Chem.* **287**, 30952–30966.

Lemaire, P. A., Tessmer, I., Craig, R., Erie, D. A. & Cole, J. L. (2006). *Biochemistry*, **45**, 9074–9084.

Li, X., Romero, P., Rani, M., Dunker, A. K. & Obradovic, Z. (1999). *Genome Inform. Ser. Workshop Genome Inform.* **10**, 30–40.

Linding, R., Jensen, L. J., Diella, F., Bork, P., Gibson, T. J. & Russell, R. B. (2003). *Structure*, **11**, 1453–1459.

Maas, C., Tagnaouti, N., Loebrich, S., Behrend, B., Lappe-Siefke, C. & Kneussel, M. (2006). *J. Cell Biol.* **172**, 441–451.

Maric, H. M., Mukherjee, J., Tretter, V., Moss, S. J. & Schindelin, H. (2011). *J. Biol. Chem.* **286**, 42105–42114.

McGuffin, L. J., Bryson, K. & Jones, D. T. (2000). *Bioinformatics*, **16**, 404–405.

Meier, J., Vannier, C., Sergé, A., Triller, A. & Choquet, D. (2001). *Nature Neurosci.* **4**, 253–260.

Mukherjee, J., Kretschmannova, K., Gouzer, G., Maric, H. M., Ramsden, S., Tretter, V., Harvey, K., Davies, P. A., Triller, A., Schindelin, H. & Moss, S. J. (2011). *J. Neurosci.* **31**, 14677–14687.

Papadopoulos, T. & Soykan, T. (2011). *Front. Cell. Neurosci.* **5**, 11.

Petoukhov, M. V., Franke, D., Shkumatov, A. V., Tria, G., Kikhney, A. G., Gajda, M., Gorba, C., Mertens, H. D. T., Konarev, P. V. & Svergun, D. I. (2012). *J. Appl. Cryst.* **45**, 342–350.

Petoukhov, M. V., Konarev, P. V., Kikhney, A. G. & Svergun, D. I. (2007). *J. Appl. Cryst.* **40**, s223–s228.

Petoukhov, M. V. & Svergun, D. I. (2005). *Biophys. J.* **89**, 1237–1250.

Pollastri, G., Przybylski, D., Rost, B. & Baldi, P. (2002). *Proteins*, **47**, 228–235.

Porod, G. (1982). *Small-Angle X-ray Scattering*. London: Academic Press.
Ratcliff, G. C. & Erie, D. A. (2001). *J. Am. Chem. Soc.* **123**, 5632–5635.
Roessle, M. W., Klaering, R., Ristau, U., Robrahn, B., Jahn, D., Gehrmann, T., Konarev, P., Round, A., Fiedler, S., Hermes, C. & Svergun, D. (2007). *J. Appl. Cryst.* **40**, s190–s194.
Roth, H. M., Römer, J., Grundler, V., Van Houten, B., Kisker, C. & Tessmer, I. (2012). *DNA Repair*, **11**, 286–293.
Round, A. R., Franke, D., Moritz, S., Huchler, R., Fritsche, M., Malthan, D., Klaering, R., Svergun, D. I. & Roessle, M. (2008). *J. Appl. Cryst.* **41**, 913–917.
Sacho, E. J., Kadyrov, F. A., Modrich, P., Kunkel, T. A. & Erie, D. A. (2008). *Mol. Cell*, **29**, 112–121.
Saiyed, T., Paarmann, I., Schmitt, B., Haeger, S., Sola, M., Schmalzing, G., Weissenhorn, W. & Betz, H. (2007). *J. Biol. Chem.* **282**, 5625–5632.
Schrader, N., Kim, E. Y., Winking, J., Paulukat, J., Schindelin, H. & Schwarz, G. (2004). *J. Biol. Chem.* **279**, 18733–18741.
Schwarz, G., Mendel, R. R. & Ribbe, M. W. (2009). *Nature (London)*, **460**, 839–847.
Schwarz, G., Schrader, N., Mendel, R. R., Hecht, H.-J. & Schindelin, H. (2001). *J. Mol. Biol.* **312**, 405–418.
Shkumatov, A. V., Chinnathambi, S., Mandelkow, E. & Svergun, D. I. (2011). *Proteins*, **79**, 2122–2131.
Sola, M., Bavro, V. N., Timmins, J., Franz, T., Ricard-Blum, S., Schoehn, G., Ruigrok, R. W., Paarmann, I., Saiyed, T., O'Sullivan, G. A., Schmitt, B., Betz, H. & Weissenhorn, W. (2004). *EMBO J.* **23**, 2510–2519.
Sola, M., Kneussel, M., Heck, I. S., Betz, H. & Weissenhorn, W. (2001). *J. Biol. Chem.* **276**, 25294–25301.
Svergun, D. I. (1992). *J. Appl. Cryst.* **25**, 495–503.
Svergun, D. I. (1999). *Biophys. J.* **76**, 2879–2886.
Svergun, D., Barberato, C. & Koch, M. H. J. (1995). *J. Appl. Cryst.* **28**, 768–773.
Svergun, D. I., Petoukhov, M. V. & Koch, M. H. J. (2001). *Biophys. J.* **80**, 2946–2953.
Tretter, V., Jacob, T. C., Mukherjee, J., Fritschy, J. M., Pangalos, M. N. & Moss, S. J. (2008). *J. Neurosci.* **28**, 1356–1365.
Tretter, V., Kerschner, B., Milenkovic, I., Ramsden, S. L., Ramerstorfer, J., Saiepour, L., Maric, H. M., Moss, S. J., Schindelin, H., Harvey, R. J., Sieghart, W. & Harvey, K. (2011). *J. Biol. Chem.* **286**, 37702–37711.
Tretter, V., Mukherjee, J., Maric, H. M., Schindelin, H., Sieghart, W. & Moss, S. J. (2012). *Front. Cell. Neurosci.* **6**, 23.
Tria, G., Kachala, M. & Svergun, D. I. (2013). *Proceedings of the 15th International Small-Angle Scattering Conference (SAS2012)*, edited by D. J. McGillivray, J. Trewhella, E. P. Gilbert & T. L. Hanley, abstract 00102. Sydney: ANSTO.
Tyagarajan, S. K., Ghosh, H., Harvey, K. & Fritschy, J. M. (2011). *J. Cell Sci.* **124**, 2786–2796.
Tyagarajan, S. K., Ghosh, H., Yevenes, G. E., Nikonenko, I., Ebeling, C., Schwerdel, C., Sidler, C., Zeilhofer, H. U., Gerrits, B., Muller, D. & Fritschy, J. M. (2011). *Proc. Natl Acad. Sci. USA*, **108**, 379–384.
Vacic, V., Uversky, V. N., Dunker, A. K. & Lonardi, S. (2007). *BMC Bioinformatics*, **8**, 211.
Volkov, V. V., Lapuk, V. A., Kayushina, R. L., Shtykova, E. V., Varlamova, E. Yu., Malfois, M. & Svergun, D. I. (2003). *J. Appl. Cryst.* **36**, 503–508.
Volkov, V. V. & Svergun, D. I. (2003). *J. Appl. Cryst.* **36**, 860–864.
Ward, J. J., McGuffin, L. J., Bryson, K., Buxton, B. F. & Jones, D. T. (2004). *Bioinformatics*, **20**, 2138–2139.
Winzer, A. T., Kraft, C., Bhushan, S., Stepanenko, V. & Tessmer, I. (2012). *Ultramicroscopy*, **121**, 8–15.
Xiang, S., Nichols, J., Rajagopalan, K. V. & Schindelin, H. (2001). *Structure*, **9**, 299–310.
Zita, M. M., Marchionni, I., Bottos, E., Righi, M., Del Sal, G., Cherubini, E. & Zacchi, P. (2007). *EMBO J.* **26**, 1761–1771.