Synaptic Arrangement of the Neuroligin/β-Neurexin Complex Revealed by X-Ray and Neutron Scattering

Davide Comoletti,1,6,* Alexander Grishaev,2,6 Andrew E. Whitten,3,4 Igor Tsigelny,1 Palmer Taylor,1 and Jill Trewhella4,5

1 Department of Pharmacology, Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California, San Diego, La Jolla, CA 92093, USA
2 National Institute of Diabetes and Digestive and Kidney Diseases, Bethesda, MD 20892, USA
3 Bragg Institute, Australian Nuclear Science and Technology Organization, Menai, New South Wales 2234, Australia
4 School of Molecular and Microbial Biosciences, University of Sydney, New South Wales 2006, Australia
5 Department of Chemistry, University of Utah, Salt Lake City, UT 84112, USA
6 These authors contributed equally to this work.

*Correspondence: dcomolet@ucsd.edu
DOI 10.1016/j.str.2007.04.010

SUMMARY

Neuroligins are postsynaptic cell-adhesion proteins that associate with their presynaptic partners, the neurexins. Using small-angle X-ray scattering, we determined the shapes of the extracellular region of several neuroligin isoforms in solution. We conclude that the neuroligins dimerize via the characteristic four-helix bundle observed in cholinesterases, and that the connecting sequence between the globular lobes of the dimer and the cell membrane is elongated, projecting away from the dimer interface. X-ray scattering and neutron contrast variation data show that two neurexin monomers, separated by 107 Å, bind at symmetric locations on opposite sides of the long axis of the neuroligin dimer. Using these data, we developed structural models that delineate the spatial arrangements of different neuroligin domains and their partnering molecules. As mutations of neurexin and neuroligin genes appear to be linked to autism, these models provide a structural framework for understanding altered recognition by these proteins in neurodevelopmental disorders.

INTRODUCTION

Synaptic connectivity is tightly regulated during development in the central nervous system. Owing to the asymmetric nature of the synaptic connections, some of the molecules involved in partnering within the pre- and postsynaptic regions of the synapse have heterophilic recognition and adhesive properties (Ferreira and Paganoni, 2002). Neuroligins and neurexins show such properties; in fact, in vitro overexpression or RNAi knockdown of these respective post- and presynaptic proteins modulates clustering of associated synaptic proteins, suggesting a critical role in forming and/or maintaining synapses (Scheiffele et al., 2000; Graf et al., 2004; Chih et al., 2005). In vivo, neuroligin function appears to be more critical for synapse maturation than for the initial formation of synaptic contacts (Varoqueaux et al., 2006). Polymorphisms of the coding regions of neurexin and neuroligin genes (including point mutations, truncations, and exon deletions) were recently found to be associated with autism spectrum disorders and mental retardation, indicating a strong genetic link to neurodevelopmental disorders (Jamain et al., 2003; Laumonnier et al., 2004; Yan et al., 2005; Talebizadeh et al., 2006; Feng et al., 2006; Szatmari et al., 2007).

The neuroligins comprise a family of transmembrane proteins composed of an extracellular, N-linked glycosylated domain with strong sequence homology to acetylcholinesterase (AChE), a Ser-Thr-rich stalk domain that carries both N- and O-linked oligosaccharides (Ichtchenko et al., 1996; Bolliger et al., 2001; Hoffman et al., 2004), a single transmembrane domain, and a small, intracellular C-terminal domain. The neuroligins bind in vitro to both α- and β-neurexin (NXα and NXβ) in a Ca2+-dependent manner (Ichtchenko et al., 1995; Boucard et al., 2005). A Ca2+-binding site has been identified in NXβ, but the role of Ca2+ binding in the association and its physiological functions are unclear.

The selectivity of neuroligin/neurexin association is tightly regulated by a hierarchy of structural determinants. First, different neuroligin isoforms (NL1–4) are encoded by separate genes and have different affinities for NXβ (Comoletti et al., 2006; Graf et al., 2006), suggesting that expression of individual neuroligin isoforms is associated with discrete synaptic pathways. Second, alternative splicing of mRNAs encoding both neurexins and neuroligins provides a basis for additional discrimination via the multiplicity of potential gene products (Graf et al., 2006; Chih et al., 2006; Comoletti et al., 2006). Finally, glycosylation and its processing of an alternatively spliced region in NL1 negatively regulate neurexin binding (Comoletti et al., 2003; Boucard et al., 2005), thus offering posttranslational control of selectivity of association.
Whereas the crystal structures of the second and sixth LNS (laminin/neurexin/SHBG-like) domains of NXα have been solved (Rudenko et al., 1999; Scheckler et al., 2006), only models based on the domain homology with AChE are currently available for the neuroligins. An initial monomeric homology model (Tsigelny et al., 2000) was refined by using experimental data on the cysteine connectivity and the sites of O- and N-linked oligosaccharide attachment of the extracellular domain of NL1 (Hoffman et al., 2004). This study confirmed that the neuroligins belong to the α/β-hydrolase fold family of proteins, a common fold shared by the cholinesterases and several other serine hydrolases. However, two additional N-linked glycosylation sites, two positions of alternative splicing, and the linkage of the disulfide bond between Cys512 and Cys546 in NL1 reveal complexity in the NL1 structure beyond the predictions based solely on AChE homology. Although the influence of splice inserts A and B on the recognition of neuroligin for NXβ has been widely investigated (Boucard et al., 2005; Chih et al., 2006; Comoletti et al., 2006; Graf et al., 2006), their locations with respect to the interacting neuroligin/neurexin surface are not known. Using analytical ultracentrifugation data, the extracellular domains of the neuroligins were shown to exist as stable dimers (Comoletti et al., 2003, 2006), and a dimeric model based on the crystal structure of the dimeric AChE (Sussman et al., 1991; Bourne et al., 1995) was proposed (Dean et al., 2003). However, definitive data on the dimerization surface of the neuroligins and on the tertiary structure and orientation of the stalk domain with respect to the globular domain of the neuroligins and the postsynaptic membrane are lacking.

We report here the results of small-angle solution scattering experiments on the extracellular domains of the neuroligins and the complex formed between NL1 and NXβ. X-ray scattering data provide overall shape information on the extracellular domains of the neuroligins and the initial structural definition of the O-linked glycosylated domain linking the extracellular domain to the transmembrane span. We also use the X-ray scattering data to refine our homology model for the NL1 dimer. A combination of X-ray scattering and neutron contrast variation data with ab initio and rigid-body modeling has yielded a structural model of the extracellular domain of NL1 complexed with NXβ. Our models are presented in relation to the synaptic disposition of the complex, and, while of inherently low resolution, they provide an important structural framework for linking genetic information on mutated neurexins and neuroligins with neurodevelopmental disorders.

RESULTS

Protein Sample Quality for Small-Angle Scattering

Extraction of structural information on individual protein molecules from solution scattering data requires that all protein molecules in solution are identical (i.e., pure, monodisperse solutions) and sufficiently dilute such that there are no distance correlations between molecules. Distance correlations between molecules (usually due to Coulombic repulsion) give rise to interparticle interference effects that suppress the small-angle scattering data and result in artificially small measured Rg values. These interparticle interference effects are linear with concentration and can be removed by extrapolation to infinite dilution. Hence, accurate structural analyses require data extrapolated to infinite dilution or measured at low enough concentrations to ensure that the interparticle interference effects are negligible.

The purity of our protein samples was established by SDS-PAGE and analytical gel filtration (Figures S1A and S1B; see the Supplemental Data available with this article online), and the monodispersity was verified by the scattering data. The observed decrease in Rg values for the NL1-638-Δ(A&B) and NL1/NXβ complex (Table 1) as concentration is increased is evidence that the samples are monodisperse and that there is interparticle interference. Scattering profiles, I(Q) versus Q, measured at the low-concentration limit, where interparticle interferences effects are negligible, for each of the neuroligins are shown for comparison in Figure 1B. Additional evidence of sample monodispersity comes from the linearity of the Guinier plots within the Guinier region (QRg < 1.3) (Figure 1B, inset), and the volumes determined for each neuroligin dimer by using the Porod invariant. These volumes show a close correspondence to the values estimated from the partial-specific volume and the NL1 dimer molecular weights determined by mass spectrometry (Comoletti et al., 2006).

Analysis of the Scattering Data of NL1 Is Consistent with the Expected Dimer Structure and Indicates that the Stalk Domain Is Extended

The P(r) profile represents the probable distribution of interatomic distances within the molecule and provides some insight into the molecular structure independent of model structures. The P(r) profile for the extracellular domain of NL1 truncated after the α/β-hydrolase fold domain at residue 638 (NL1-638) gives values for Rg and Dmax of 42.7 ± 0.7 Å and ~130 Å, respectively (Figure 2A; Table 1). These values are consistent with previous hydrodynamic determinations of NL1 as a dimer (Comoletti et al., 2003, 2006). NL1-691 represents the entire extracellular domain of NL1, including a S3 amino acid segment carrying several O-linked oligosaccharides and an N-linked glycosylation site at position N662 (Hoffman et al., 2004). The presence of oligosaccharides, combined with an abundance of Pro residues, presumably maintains the peptide chain in a semirigid, bottlebrush-like structure, as demonstrated in the P-selectin molecule and the T cell coreceptor CD8 protein (Li et al., 1996; Merry et al., 2003). Inspection of the P(r) profile of NL1-691, compared to NL-638 (Figure 2A), reveals that the main peak broadens and moves to larger r values: Dmax increases by ~35 Å, and a low frequency of distances appears beyond ~130 Å, approximately where the P(r) function of the shorter NL-638 mutant approaches zero. These changes indicate that the stalk domain extends away from the main globular domain (Figure 2A, inset). The prevalence of both N- and O-linked oligosaccharides in the C-terminal sequence is
in accord with the larger volume of NL1-691 compared to NL1-638 (Table 1).

NL1-638-Δ(A&B) is the reference NL1 structure used for all of our modeling. This NL1 variant, which lacks both splice inserts A and B, which are present in NL-638, gives values for \( R_g \) and \( D_{\text{max}} \) of 42.5 ± 0.4 Å and ~130 Å, respectively (Figure 2B: Table 1, Stanford Synchrotron Radiation Laboratory [SSRL] data). The agreement between the low concentration measurement for NL-638-Δ(A&B) taken at SSRL and data from the same sample preparation taken by using the University of Utah instrument and extrapolated to infinite dilution (Table 1) verifies that SSRL samples are a good approximation to infinite dilution. The \( P(r) \) function determined for NL-638-Δ(A&B) extends to the same \( D_{\text{max}} \) value determined for NL1-638, indicating that the splice inserts do not affect the longest distances in the structure. There is a small increase in the frequency of vector lengths between ~50 and 90 Å for NL-638, consistent with the presence of the two additional loops folded onto each globular domain.

### Table 1. Scattering Data Statistics of the Neuroligins and the Neuroligin/Neurexin Complex

<table>
<thead>
<tr>
<th>Sample</th>
<th>Protein Conc.(^a) (mg/ml)</th>
<th>( R_g ) (Å)</th>
<th>( D_{\text{max}} ) (Å)</th>
<th>Porod Volume(^b) (10(^3) × Å(^3))</th>
<th>Calculated Volume(^c) (10(^3) × Å(^3))</th>
<th>MW(^d) (kDa)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>SSRL Data</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NL1-638-Δ(A&amp;B)</td>
<td>1.8</td>
<td>42.5 ± 0.4</td>
<td>130</td>
<td>209 ± 20</td>
<td>198</td>
<td>130/151/144</td>
</tr>
<tr>
<td>NL1/NX complex</td>
<td>3.6</td>
<td>46.8 ± 0.2</td>
<td>155</td>
<td>*</td>
<td>275</td>
<td>181/201/199</td>
</tr>
<tr>
<td><strong>University of Utah Instrument Data</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NL1-638-Δ(A&amp;B)</td>
<td>13.2</td>
<td>38.3 ± 0.3</td>
<td></td>
<td></td>
<td></td>
<td>130/151/144</td>
</tr>
<tr>
<td></td>
<td>7.6</td>
<td>40.1 ± 0.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>4.1</td>
<td>41.4 ± 0.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inf. dilution</td>
<td>42.4 ± 0.6</td>
<td>130</td>
<td>208 ± 14</td>
<td>198</td>
<td></td>
<td>130/151/144</td>
</tr>
<tr>
<td>NL1-638</td>
<td>3.3</td>
<td>42.7 ± 0.7</td>
<td>130</td>
<td>250 ± 19</td>
<td>220</td>
<td>136/166/160</td>
</tr>
<tr>
<td>NL1-691</td>
<td>3.8</td>
<td>51.8 ± 1.0</td>
<td>165</td>
<td>255 ± 26</td>
<td>257</td>
<td>148/189/185</td>
</tr>
<tr>
<td>NL2-615</td>
<td>3.7</td>
<td>40.6 ± 0.6</td>
<td>130</td>
<td>178 ± 7</td>
<td>193</td>
<td>135/146/140</td>
</tr>
<tr>
<td>NL3-639</td>
<td>1.2</td>
<td>40.3 ± 0.7</td>
<td>130</td>
<td>164 ± 12</td>
<td>190</td>
<td>128/144/138</td>
</tr>
<tr>
<td>NL4-619</td>
<td>3.4</td>
<td>42.1 ± 0.6</td>
<td>135</td>
<td>199 ± 7</td>
<td>200</td>
<td>132/140/145</td>
</tr>
<tr>
<td>NL1/NX complex</td>
<td>19.7</td>
<td>40.9 ± 0.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>15.7</td>
<td>40.8 ± 0.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>9.8</td>
<td>44.0 ± 0.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>8.7</td>
<td>43.7 ± 0.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>6.6</td>
<td>44.5 ± 0.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>4.5</td>
<td>45.2 ± 0.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3.8</td>
<td>47.7 ± 0.9</td>
<td>155</td>
<td>*</td>
<td>275</td>
<td>181/201/199</td>
</tr>
<tr>
<td>Inf. dilution</td>
<td>47.7 ± 0.8</td>
<td>155</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Neutron Data, 42% D(_2)O</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NL1/deuterated NX complex</td>
<td>5.1</td>
<td>49.0 ± 2.0</td>
<td>155</td>
<td>n.a.</td>
<td>n.a.</td>
<td>130/151/144</td>
</tr>
</tbody>
</table>

\(^a\) Concentrations for all of the NL constructs (except NL-691) are determined from I(0) analysis by comparison to a lysozyme standard measured in the same cell on the same day. The NL-691 concentration is from a Bradford assay. Concentrations for the complexes are all based on quantitative amino acid analysis. Infinite dilution refers to the results from the scattering data concentration series extrapolated to zero concentration.

\(^b\) Volumes calculated from the scattering data by using the Porod invariant (Porod, 1951). The asterisk indicates that the volumes for these data sets varied significantly with the Q max value chosen; thus, the uncertainty of these values is much larger than the propagated statistical errors quoted for the other volumes and are not considered to be reliable.

\(^c\) Calculated from the molecular weight (Da)/partial specific volume (v) where MW was determined from mass spectrometry measurements.

\(^d\) These molecular weight values were calculated as follows: from peptide sequence; from peptide sequence plus assuming full glycosylation; from mass spectrometry data on independent samples.
Figure 1. Schematic Representation of the Extracellular Domains of the Neuroligins and NX$i$ and Scattering Profiles and Guinier Plots of the Neuroligins

(A) Domain organization of the full-length NL1 is shown on top. Different recombinant neuroligin isoforms were generated by truncation of the AChE-homologous domain, as shown by the numbers on the right side of each domain. The numbering on top of the full-length NL1 refers to residues flanking splice inserts A and B. The amino acid numbering of NL1–3 is according to Ichtchenko et al. (1996); that for NL4 is according to Bolliger et al. (2001). The numbers on the NX$i$ construct refer to the mutations described in the text and are in accord with Ushkaryov et al. (1994).
Neuroligins 1–4 Form Dimers of Similar Dimensions with the Dimerization Domains Resembling Those of AChE

As expected from the high sequence identities (72%–77%) among the neuroligins (Ichtchenko et al., 1996; Bolliger et al., 2001), our scattering data and associated parameters, such as the $D_{\text{max}}$ and $R_g$, values show that the extracellular domains of all of the neuroligins are similar in overall shape and dimensions, indicating that they share a common fold and have similar dimerization domains (Table 1). Inspection of the $P(r)$ functions of the soluble extracellular domains of NL1–4 (Figure 2C) reveals only minor deviations from our NL1 reference structure. Physical parameters and $P(r)$ profiles from scattering data collected on mouse AChE truncated at position 548 (Marchot et al., 1996) indicate that the NL structures are also very similar to the AChE structure, supporting our homology model as a reliable starting point for the solution structure of all of the neuroligins (Figure S2).

The starting model for rigid-body refinement against the NL1-638-D(A&B) data was obtained by manually docking two NL1 monomers, by using the AChE crystal structure as a guide. The crystal structure of Torpedo californica AChE shows two monomeric units connected by a four-helix bundle containing a cysteine bond, thus forming a covalent dimer (Sussman et al., 1991). The same dimerization motif was found in the crystal structures of mouse AChE lacking that C-terminal cysteine, indicating that dimerization of the AChEs does not require a covalent bond (Bourne et al., 1995). For our NL1-638-D(A&B) model, residues 450–462 and residues 621–634 made up the two helices that form the dimerization domain. This assembly was then optimized by simulated annealing/molecular dynamics structure refinement of the relative positions of the monomer units against X-ray scattering data (Figure 3). When compared to our initial homology model, the refined structure had one of the monomers rotated ~20° about the long axis of the dimer molecule, and, as a result, the atomic coordinates of the backbone atoms have ~6 Å root mean square distance displacement with respect to the initial coordinates (Figure 3).

Shape restoration, using ab initio methods, provides an alternate modeling approach to the interpretation of scattering data that has minimal assumptions about starting geometries. We therefore used ab initio shape restoration (implemented in the programs DAMMIN, DAMAVER, and DAMFILT) to derive the most probable shape that best fit the scattering data for NL1. Figure 3 shows the excellent agreement between this most probable shape and the atomic model derived from the rigid-body refinement. Similar results were obtained from shape-restoration calculations of AChE-548 (Figure S2) and for neuroligins 2–4 (Figure S3). The $P(r)$ function for NL1-691 exhibits significantly increased $R_g$ and $D_{\text{max}}$ values, which likely result from the C-terminal Ser-Thr-rich domain extension.
However, this feature was not resolved in the shape-restoration calculations, possibly due to its relatively narrow, single-chain structure and its likely segmental motion in the absence of cell membrane tethering.

Two NX$\beta$ Molecules Bind on Opposite Sides of the Long Axis of the NL1 Dimer in a Defined Orientation and with a Median Separation of 107 Å

As NL1-638-Δ(A&B) has the highest affinity for NX$\beta$ (Comoletti et al., 2006), it was chosen to study the complex. To ensure that measurements were performed on fully complexed samples (i.e., 2:2 stoichiometry), protein concentrations were determined by quantitative amino acid analysis and were checked by I(0) analysis after data collection. To improve the statistical precision of the data used for modeling, we took the average of three independent X-ray scattering data sets measured at the low-protein concentration limit determined to exclude interparticle interference (Figures 4A and 4B). The corresponding P(r) profile yielded $R_g$ and $D_{max}$ values of 47.7 ± 0.9 Å and ~155 Å, respectively (Table 1). The ~25 Å increase in $D_{max}$ compared to the free NL1-638-Δ(A&B), accompanied by a small shift of the main peak (~40–45 Å) and an increase in intensity between ~60 Å and ~110 Å (Figure 4C, triangles), indicates that NX$\beta$ binds at the periphery of the NL1 dimer, distant from the dimer symmetry axis.

Ab initio shape restoration from the scattering data for the complex resulted in ~50% of the DAMMIN calculations converging to a single class of shapes that shows two clearly recognizable protrusions from the basic NL1 dimer shape in locations that are consistent with the binding data and with the NX$\beta$ separation distance obtained from the neutron contrast variation experiment (below). The remaining 50% of the conformations showed degenerate solutions that were inconsistent with these data. As this same degeneracy was observed when modeling theoretical intensity profiles calculated by using the program CRYSOL and the atomic model we derive for our complex (below), it appears that the degenerate solutions are artifacts of the reconstruction algorithm for this shape.

In order to determine unambiguously the positions of the NX$\beta$ subunits relative to the NL1 dimer, we required an experiment that could distinguish between the NL1 and NX$\beta$ components. The signal intensity in a small-angle scattering experiment is proportional to the square of the difference between the scattering densities of the particle and the solvent (the “contrast”). As isotopes of hydrogen ($^{1}$H = H) and deuterium ($^{2}$H = D) have different neutron “scattering powers,” the neutron scattering density of a protein, as well as its solvent, can be systematically varied by substituting H for D. Therefore, deuterated NX$\beta$ was prepared, and stoichiometric amounts of deuterated NX$\beta$ were combined with NL1-638-Δ(A&B) in the presence of 2 mM Ca$^{2+}$. By acquiring neutron scattering data from the complex in a 42% D$_2$O solution, the protonated NL1 subunits were effectively “solvent matched”; that is, they had the same approximate mean neutron scattering density as the solvent and thus did not contribute significantly to the scattering signal, allowing for direct measurement of the scattering from the NX$\beta$ units. The $P(r)$ curve extracted

![Figure 3. Ribbon Representation of the Dimeric NL1 Model Superimposed on the Ab Initio Shape-Restoration Envelope Derived from the Scattering Data](image-url)
from the neutron scattering data (Figure 4C, open circles) exhibits two maxima at ~20 Å and 107 Å: the first corresponds to the most frequent interatomic distance within the individual NXβ units, and the second corresponds to the most frequent interatomic distance between the two NXβ subunits. The combined X-ray and neutron scattering data thus provide the approximate binding locations for the NXβ subunits and the separation of their centers of mass: a powerful set of restraints for refining a solution structure of the complex.

The initial starting point for the rigid-body modeling procedure was obtained by placing the NXβ subunits proximal to the NL1-638-Δ(A&B) dimer to best fit the low-resolution density reconstruction of the complex and according to known biochemical constraints (Comoletti et al., 2003, 2006; Boucard et al., 2005; Graf et al., 2006). A series of rotations of the NXβ subunits within this model allowed us to sample a diverse set of initial geometries for the rigid domain refinement. Refinement against X-ray data from all starting geometries converged to only two alternate models related by the 180° rotation of the NXβ monomers, both of which had the rim of the β sheet sandwich opposite the N and C termini of NXβ as the site of contact with NL1. The scattering data thus proved to be sensitive to the orientation of the discoid shape of the NXβ subunits, allowing us to select the best fitting of the two for additional refinement against the X-ray scattering data while restraining the NXβ/NXβ inter-domain separation according to the neutron scattering data. This final calculation produced a structure in which the NXβ monomers are positioned in agreement with the available mutagenesis data (Comoletti et al., 2006; Graf et al., 2006) and the C termini of the NL1 and NXβ subunits facing opposite directions, consistent with the complex bridging pre- and postsynaptic surfaces (Figure 5).

**DISCUSSION**

Using small-angle solution scattering, we determined the three-dimensional structural model of the AChE-like domain of the neuroligins, showing that all four neuroligins have similar overall shapes and dimensions, as evidenced by their $P(r)$ functions, $R_g$ and $D_{max}$ values, and three-dimensional shape reconstructions. The stalk region connecting the globular domain of the neuroligins with their transmembrane domain is elongated and projects away from the globular domain. This fundamental information on the neuroligin structure, based on X-ray scattering and neutron contrast variation experiments, enables us to present the first, to our knowledge, three-dimensional structural model of NL1 in a complex with NXβ and to place it in the context of the synapse.

The very good fit of X-ray scattering data with the neuroligin model, the similarity with the solution structure determined for the AChE whose crystal structure has been extensively studied, and the correspondence among their physical parameters strongly indicate that the neuroligin dimer is held together by the four-helix bundle typical of the $\alpha/\beta$-hydrolase fold (Bourne et al., 1995). This arrangement of the dimerization domain results in each stalk...
region emerging from the same side of the dimer. The very small differences in distance distributions \((P(r))\) between the spliced and unspliced forms of NL1 indicates that splice inserts A and B in the NL1 fold are within the main globular domain, consistent with their role in neurexin association (Boucard et al., 2005; Comoletti et al., 2006).

A puzzling element of the structure of these heterophilic adhesion proteins is the relatively small mass of their extracellular domains when compared with other transmembrane synaptic proteins such as N-cadherins, L1, and NCAM (Piechotta et al., 2006). As such, the globular extracellular domains of the neurexins and neuroligins, even when associated as a complex, are too small to span the synaptic space. However, the O-linked glycosylated sequence with its relatively rigid, bottle-brush-like structure offers an extended stalk region intervening between the AChE-homologous domain and membrane-spanning regions. This structure may only be observable in solution studies, as filamentous and glycosylated domains generally preclude crystallization. As the neurexins also have a similar stalk region between their LNS and transmembrane domains, it is expected that both stalk regions act as intervening lines of variable extension linking the sequence emerging from the membrane to the globular domains. Hence, the trans-synaptic space can be spanned from presynaptic and postsynaptic membrane tethers.

In LNS domains, the metal-binding pocket is located at the rim of the \(\beta\) sheet sandwich opposite the N and C termini (Rudenko et al., 2001). In N\(\xi\), alanine mutations of D137 and N238, residing on this rim, lack synaptogenic activity (Graf et al., 2006), possibly participating in the binding of \(Ca^{2+}\). On the NL1 surface, mapping N\(\xi\)-binding determinants revealed that the region flanking splice insert B (including E297 and K306) (Boucard et al., 2005; Comoletti et al., 2006) is the focal point for N\(\xi\) binding, consistent with our solution scattering data on the NL1/N\(\xi\) complex. In addition, the neuroligin surface near splice insert B is well removed from the other three sites of glycosylation, that were shown to have no influence on neurexin association (Comoletti et al., 2003).

Based on secondary-structure predictions, it was suggested that splice insert 4 of neurexin forms two helices located in the long loop connecting strands \(\beta\)10 and \(\beta\)11 (Rudenko et al., 1999). Although initial reports based on in vitro pull-down and cell-associated studies described splice insert 4 as the main switch for NL1 binding (Ichtchenko et al., 1995, 1996), N\(\xi\) interaction with the neuroligins is modulated in a more complex manner (Boucard et al., 2005; Graf et al., 2006; Comoletti et al., 2006). Consistent with our results, the insertion point of splice insert 4 in the LNS domain does not directly contact the NL1/N\(\xi\) interface.

In the synapse, orientation of the neuroligins and neurexins in the complex requires that their C-terminal regions extend the O-linked glycosylated domains in opposite directions. Since both N and C termini emerging from the LNS domain were disordered in the crystal structure (Rudenko et al., 1999), establishing an orientation a priori of the C-terminal end of the N\(\xi\) monomer was not possible. Nevertheless, our rigid-body structure refinement of X-ray and neutron solution scattering data favors one particular orientation of the N\(\xi\) monomers in which interfacial residues of NL1 and N\(\xi\) are defined by the available mutagenesis data. In a recently published model (Dean and Dresbach, 2006), it is proposed that two neurexin monomers bind to the apical ends of the two neuroligin monomers, an arrangement inconsistent with the available
biochemical data, as well as the X-ray and neutron scattering data reported in this study. Synaptic-differentiation studies suggest that a minimum of four neurexin molecules have to associate to obtain presynaptic activation in vitro (Dean et al., 2003), whereas solution studies clearly demonstrate that a stable 2:2 complex is formed (Comolatti et al., 2006). These results imply that formation of larger clusters of these synaptic proteins may occur through intracellular domains.

An R451C mutation of the NL3 gene and three mutations of the NL4 gene (G99S, K378R, and V403M) were recently found in autistic patients (Jamain et al., 2003; Yan et al., 2005). These amino acids are highly conserved in the neuroligin family and are located in the AChE-like domain. Upon mapping these mutations in our structural model of the NL1/NXβ complex, it is interesting to note that NL4-K378R, NL4-V403M, and NL3-R451C cluster in a region opposite the NXβ-binding region on each monomer and relatively close to the symmetry axis of the dimer. In contrast, NL4-G99S is found on a surface loop at the extreme end of the long axis of the neuroligin dimer (Figure 6). Owing to their positions with respect to the NXβ-binding
interface, we predict that these mutations do not directly influence neuroligin/neurexin association. Although the R451C shows reduced NXβ binding (Comoletti et al., 2004), it was recently shown that the mutation restricts processing through the endoplasmic reticulum probably because of local protein misfolding and an inability to oligomerize (De Jaco et al., 2006).

**Conclusions**

Glycosylation of the neuroligins and the extended stalk region has rendered a crystallographic solution of the structure of the extracellular portion of the neuroligins and their complex with NXβ a challenging endeavor. Thus, we turned to solution scattering to determine the molecular shapes of the individual molecules in order to enhance our understanding of the complex formation within the synaptic space environment. Accordingly, solution scattering obviates the constraints dictated by the crystal packing (Heidorn and Trewhella, 1998). The presented structural models for NL1 and NL1/NXβ are of inherently low resolution and are based on initial AChE homology as a reasonable starting point for the neuroligin dimer, for which there is good supporting evidence. Our study thus has afforded additional analysis of unstructured neuroligins, and the model of the neurexin and neuroligin associated in the synaptic space (Figure 7) provides a structural framework for analyzing the disposition of these molecules within the synaptic space.

**EXPERIMENTAL PROCEDURES**

**Expression Vectors and Mutagenesis**

FLAG-NL1–NL4 cDNAs, devoid of both transmembrane and intracellular domains, were used to make the proteins NL1-638, NL2-615, NL3-639, and NL4-619, as described previously (Comoletti et al., 2006). NL1-638 contained both splice inserts A and B, whereas NL2-615, NL3-639, and NL4-619 did not contain either splice insert. NL1-638 was also engineered devoid of splice inserts A and B (NL1-638-Δ(A&B)). Soluble NL1-691 containing both splice inserts A and B and the O-linked glycosylated stalk domain was prepared by introducing a stop codon at Tyr692 (Comoletti et al., 2003). All mutations were obtained by using the QuikChange Mutagenesis Kit (Stratagene, San Diego, CA) and were verified by both restriction digests and DNA sequencing. The GST-cleavable β-neurexin construct was described elsewhere (Comoletti et al., 2004, 2006). Soluble mouse AChE (AChE-548) was constructed by introducing a stop codon at Cys549 of mouse AChE-GPI as described previously (Marchot, et al., 1996).

**Neuroligin and Neurexin Expression and Purification**

To purify each soluble neuroligin, tissue culture medium containing the secreted protein was passed over an M2 anti-FLAG affinity column (Sigma) (Comoletti et al., 2003). Samples were concentrated by microfiltration with Microcon 30YM (Millipore) and were separated by analytical gel filtration with a Superdex 200 HR 10/30 column (GE Healthcare). SDS-PAGE gels stained with Simply Blue Safestain

![Figure 7. Model of the NL1/NXβ Assembly in the Synapse](image-url)
Structure of the Neuroligin/Neurexin Complex

(Neurogenin, Carlsbad, CA) were used to purify for clarity and the absence of degradation products. For neutron scattering experiments, deuterated NX(I) was expressed in BL21 E. coli (Invitrogen, Carlsbad, CA) by using Spectra 9 deuterated media (Spectragases, Branchburg, NJ) and was purified as described (Comoletti et al., 2006, 2004). To establish the level of neurexin deuteration, mass spectrometry on purified recombinant NX(I) was performed as described (Comoletti et al., 2006). The deuterated NX(I) had a mass of 26,085 Da, indicating that all nonexchangeable hydrogens were deuterated (Figure S1B).

Quantitative Amino Acid Analysis
Protein concentrations were estimated by amino acid composition on a Beckman 63000 analyzer after hydrolysis in 5.7 N HCl containing 0.1% phenol in vacuum, at 110°C. Only reliably quantified amino acids (Ala, Arg, Asx [Asn + Asp], Glx [Glu + Glu], Ile, Leu, Lys, Phe, and Val) were used to estimate the protein concentration of samples analyzed by small-angle scattering.

Initial Homology Models of Neuroligins
Structural models of the extracellular domain of the neuroligins were generated with the programs Homology and InsightII (Accelrys, Inc.) by using the crystal structure of mouse ACHE (PDB code: 1MAH) as a template. The structure of each neurolgin was then energy minimized for 10,000 iterations by using the distance-dependent dielectric constant with the program Discover (Accelrys, Inc.).

X-Ray Scattering Data Acquisition
X-ray scattering data for NL1-638-Δ(A&B) and its complex with NX(I) were acquired at the SSRL, beamline 4-2 (Tsuruta et al., 1998) at X-ray scattering data collection time). The first setting gives an accessible small-angle scattering. The second setting gives an accessible small-angle scattering.

Structure of the Neuroligin/Neurexin Complex

To minimize possible protein aggregation, samples that could have the effect of overfitting the data. Ten independent models were aligned and averaged by using the program DAMAVER (Volkov and Svergun, 2003). The program DAMFILT (Volkov and Svergun, 2003) was used to filter the models; the mean volume of the constituent runs was set as the cutoff to produce the average model (Figures 3 and 5). Finally, the coordinates of the homology-based dimeric model were superimposed on the filtered DAMMIN model by using the program SUPCOMB (Kozin and Svergun, 2001).

Refined atomic models of both the neurolgin dimer and the entire NL1/NX(I) complex were obtained by rigid-body molecular dynamics refinement against small-angle scattering data by using the CNS program (Brunger et al., 1998) supplemented by a small-angle X-ray scattering fitting module (Grishavaev et al., 2005). First, SSRL X-ray scattering data in the Q range between 0.022 Å⁻¹ and 0.247 Å⁻¹ were used for the NL1 dimer structure refinement. The ACHE-based homology model of the dimer was used as a starting point in the refinement; the geometries of the individual domains were fixed throughout the protocol, and the Cg symmetry of the dimer was enforced. Subsequently, X-ray data from the University of Utah instrument, in the Q range from 0.020 Å⁻¹ to 0.300 Å⁻¹, were used for further refinement; Dmax was set to 150 Å for deeming the data to remove effects of the instrument slit geometry. In all calculations for the complex, the structure of the NL1 dimer was kept identical to the refined dimer geometry obtained in the previous stage. NX(I) was fixed at the previously determined X-ray structure coordinates (PDB code: 1C4R), and the Cg symmetry of the entire complex was enforced. In order to remove any potential bias from a particular orientation of the NX(I) units within the complex, four starting geometries were generated with the NX(I) monomers rotated in 90° steps around the vector perpendicular to the NL1 surface at the site of NX(I) contact. These initial geometries had the NX(I) residues 109, 137, and 238 facing the NL1 units, in agreement with published results (Graf et al., 2006). At the final stage of refinement, the separation between the centers of mass of the neurexin subunits was restrained to 107 Å, from the SANS data acquired in 42% D2O. Final three-dimensional models were made by using PyMOL software version 0.97 (DeLano, 2004). All cited errors are based on propagated counting statistics only.

NL1-638-Δ(A&B) used in the complex characterization has three N-linked glycosylation sites. As we have previously studied the occupancy and type of carbohydrate in recombinant NL1 (Hoffman et al., 2004), this estimate is based on using the Percus-Yevick closure for the structure factor calculation as implemented in GIFT (Fritz et al., 2000).

X-Ray and Neutron Scattering Data Analysis
For an initial assessment of data quality, Guinier plots were evaluated for the expected linearity by using the program PRIMUS (Konarev et al., 2003). The pair distance distribution function, P(r), was calculated by using the indirect Fourier transform method of Svergun (1999) as implemented in the program GNOM; the P(r) function was required to go to zero at r = 0 and r = Dmax, the maximum linear dimension of the scattering particle. GNOM was also used to correct for slit geometry of the X-ray source at University of Utah. Physical parameters determined for all neuroligins and the NL1/NX(I) complex are shown in Table 1. All statistical fittings shown were judged as excellent by the fit criteria of GNOM. The fits of the final model structures to the X-ray and neutron scattering data were evaluated with CRYSOL (Svergun et al., 1995) and CRYSTAL software (Svergun et al., 1998), respectively, with corrections for the wavelength and geometry smearing as appropriate.
Structure of the Neuroligin/Neurexin Complex


Jamain, S., Quach, H., Betancur, C., Rastam, M., Colineaux, C., Gilberg, I.C., Soderstrom, H., Giros, B., Leboyer, M., and Gillberg, C.,...


