ABSTRACT: Cleavage of amyloid-β precursor protein (APP) by site-specific proteases generates amyloid-β peptides (Aβs), which are thought to induce Alzheimer’s disease. We have identified an independently folded extracellular domain of human APP localized proximal to the Aβ sequence, and determined the three-dimensional structure of this domain by NMR spectroscopy. The domain is composed of four α-helices, three of which form a tight antiparallel bundle, and constitutes the C-terminal half of the central extracellular region of APP that has been implicated in the regulation of APP cleavage. Sequence comparisons demonstrate that the domain is highly conserved among all members of the APP family, including invertebrate homologues, suggesting an important role for this region in the biological function of APP. The identification of this domain and the availability of its atomic structure will facilitate analysis of APP function and of the role of the extracellular region in the regulation of APP cleavage.
**EXPERIMENTAL PROCEDURES**

**Protein Expression and Purification.** DNA constructs for expression of the different fragments of human APP were generated by PCR with custom-designed primers, subcloned into pGEX-KG (9) or pGEX-KT (10) vectors, and expressed in bacteria (Escherichia coli BL21) as GST fusion proteins. To generate uniformly 15N-labeled or 15N- and 13C-labeled samples for NMR studies, bacteria were grown in minimal medium supplemented with 15NH4Cl or with 15NH4Cl and [13C6]glucose (CIL Inc.) as the sole nitrogen and carbon sources, respectively. Soluble fusion proteins were affinity-purified on glutathione–Sepharose beads (Pharmacia), cleaved from the GST moiety with thrombin, and further purified by ion exchange or size exclusion chromatography. The recombinant protein did not contain any detectable contaminants as judged by SDS–PAGE, the size exclusion chromatography profile, and UV and H–15N HSQC spectra.

**Limited Proteolysis.** The GST–APP recombinant fragment [residues 365–658 of human APP, produced in the pGEX-KG bacterial expression vector (9) in 15N-labeled medium] was attached to glutathione-Sepharose beads, cleaved with thrombin, and purified by ion exchange chromatography. Thrombin cleavage generates the recombinant protein bearing an additional N-terminal vector-derived sequence (GSPGISGGGGGILEV) attached to the N-terminus of the APP fragment. Aliquots of the purified recombinant protein were incubated for 30 min at room temperature with different amounts of L-1-tosylamide-2-phenylethylchloromethyl ketone (TPCK)-treated trypsin (Sigma) at the protein:enzyme ratios given in Figure 1B. Proteolytic fragments were separated by SDS–PAGE and identified by N-terminal sequencing and mass spectrometry. The observed molecular mass was 27 573 Da for the intermediate trypsin-resistant fragment and 15 000 Da for the smaller fragment (the calculated molecular mass for the 15N-labeled APP(365–585) with the N-terminal extension is 27 578 Da and for the 15N-labeled APP(461–585) is 15 004 Da).

**NMR Spectroscopy.** All NMR data were acquired on an INOVA500 spectrometer at 27 °C. Initial H–15N HSQC spectra of the APP(460–576) and APP(460–585) fragments were acquired at protein concentrations of 50–60 μM in buffer containing 25 mM sodium phosphate and 150 mM NaCl. Structure elucidation was performed using 0.85 mM uniformly 15N-labeled or 15N- and 13C-labeled samples of the APP(460–576) fragment in buffer containing 50 mM sodium phosphate, 250 mM NaCl, and 300 mM guanidinium chloride (GdnHCl) with 5% 2H2O at pH 6.4. Resonance assignments, NOEs, and amide protection data for structural determination were obtained from a series of two-dimensional (2D) and three-dimensional (3D) experiments as described previously (11). These experiments included 3D H–15N TOCSY-HSQC and NOESY-HSQC, HNCO, HNCA, CBCA(CO)NH, H(C)(CO)NH-TOCSY (21 ms mixing time), (H(C)(CO)NH)-TOCSY (21 ms mixing time), and HCCH-TOCSY (14 ms mixing time) spectra for resonance assignments (12–15), and 2D NOEY, 3D H–15N NOESY-HSQC, and 3D H–13C NOESY-HSQC spectra (all with a mixing time of 100 ms) for measuring NOEs for structure determination. Protection of amide protons from the solvent was assessed from the intensities of exchange cross-peaks with the water resonance in 3D H–15N NOESY-HSQC and TOCSY-HSQC experiments. Stereospecific assignments of valine and leucine methyl groups were obtained from a 1H–13C HSQC spectrum acquired on a 0.8 mM sample of 10% 13C-labeled APP(460–576). All data were processed with NMRPipe (16) and analyzed with NMRView (17).

**Structure Calculations.** NOE cross-peak intensities were classified as strong, medium, weak, and very weak and assigned to restraints of 1.8–2.8, 1.8–3.5, 1.8–5.0, and 1.8–6.0 Å, respectively, with appropriate pseudoatom corrections. Backbone torsion angles were derived from chemical shift analysis using TALOS (18) and employing the
standard database provided by the program. Restraints were set at 1.5 times the standard deviation yielded by TALOS (22.5° minimum). Hydrogen bond restraints were 1.7–2.5 and 2.7–3.5 Å for H–O and N–O distances, respectively. Structures were calculated by simulated annealing using torsion angle dynamics with CNS (19). The force constants used were 75 kcal mol⁻¹ Å⁻² and 400 kcal mol⁻¹ rad⁻² for distance and torsion angle restraints, respectively. A total of 1000 structures were calculated with the final set of restraints, and the 20 structures with the lowest overall energies were selected.

RESULTS

To structurally characterize the CAPPD, we performed limited proteolysis experiments with a recombinant fragment composed of residues 365–658 of human APP. This fragment, which contains the predicted CAPPD and ~90 additional residues at the C-terminal end, was chosen because sequence alignments showed that these additional residues are evolutionarily conserved in APP, but not in APLPs (Supporting Information Figure 1). We observed an intermediate trypsin-resistant fragment with a molecular mass of 30 kDa that was further proteolyzed to a smaller fragment of ~15 kDa (Figure 1B). Using mass spectroscopy and N-terminal sequencing, we found that the intermediate fragment corresponds to residues 365–585, matching the CAPPD boundaries predicted by sequence alignments. The smaller fragment arises from trypsin cleavage at two arginine residues, R460 and R585, and spans the C-terminal half of the CAPPD. Its resistance to proteolysis suggests that this fragment contains a minimal structural unit within the CAPPD that we termed CAPPD*.

To further characterize CAPPD* and better define its boundaries, we used NMR spectroscopy. Two recombinant ¹⁵N-labeled fragments encompassing CAPPD* (residues 460–585 and 460–576) were analyzed. Both recombinant proteins were monomeric under physiological conditions and at protein concentrations of <60 μM, as revealed by their one-dimensional NMR signal line widths and size exclusion chromatography profiles (data not shown). In addition, the excellent dispersion of their ¹H–¹⁵N heteronuclear single-quantum correlation (HSQC) spectra (Supporting Information Figure 2) confirmed that the fragments contain an independently folded domain. Since the ¹H–¹⁵N HSQC spectra of both fragments were practically identical and the additional cross-peaks observed for the longer fragment exhibited narrow line widths and poor ¹H chemical shift dispersion characteristic of unstructured polypeptides, we focused on the shorter fragment (residues 460–576) for structure determination. The fragment exhibited a strong tendency to...
reversibly aggregate at concentrations of >100 μM under a wide range of conditions that were tested, but became highly soluble in monomeric form at the concentrations required for structure elucidation (close to 1 mM) in the presence of 250 mM NaCl and 300 mM GdnHCl. Comparison of the structure of CAPPD* is not affected by the addition of GdnHCl.

Overhauser effects (NOEs) and the deviations of the backbone chemical shifts from random coil values showed that residues 460–569 are structured, which thus defines the boundaries of CAPPD*. Restraints derived from NOEs, chemical shifts, and amide accessibility data were then used to determine the structure of CAPPD* in solution.

A total of 1613 experimental restraints were used to generate a final set of 20 conformers of the domain. The structural statistics are summarized in Table 1, and backbone superpositions of the structures are shown in Figure 3. Ribbon diagrams and space filling models of the lowest-energy structure in two orientations are shown in Figure 4. The quality of the structures is reflected by small deviations from the experimental restraints and from idealized covalent geometry, as well as by the high percentage of residues in the most favored regions of the Ramachandran map (Table 1).

The structure of CAPPD* reveals that the domain is composed of four α-helices connected by short loops and/or turns (Figure 4A,B), in agreement with the secondary structure predicted from the NOE patterns and the backbone chemical shifts. The three N-terminal helices (HA, HB, and HC) form a tight elongated up-and-down bundle, while the fourth helix (HD) “crosses” the top of the bundle. Helix HA is significantly shorter than the other helices of the bundle, and as a result, there appears to be some “wobbling” of the bottom parts of helices HB and HC with respect to the rest of the domain. This wobbling became apparent when we superimposed the “top” and “bottom” parts of the CAPPD* structure predicted from the NOE patterns and the backbone r.m.s. deviations (0.57 ± 0.18 Å). In panel B, residues constituting the bottom part of the domain (504–537) was superimposed, resulting in a backbone r.m.s. deviation of 0.47 ± 0.15 Å.

FIGURE 3: Backbone superpositions of the 20 structures of human CAPPD* with the lowest overall energies. Since helix HA of the three-helix bundle is significantly shorter than the other two helices, the relative orientation of the bottom of helices HB and HC with respect to the rest of the domain is poorly defined. Thus, two different backbone superpositions were generated. In panel A, the top part of the domain (residues 463–504 and 537–566) was superimposed, yielding a backbone r.m.s. deviation of 0.57 ± 0.18 Å. In panel B, residues constituting the bottom part of the domain (504–537) was superimposed, resulting in a backbone r.m.s. deviation of 0.47 ± 0.15 Å.

Table 1: Structural Statistics and Atomic r.m.s. Deviations for the 20 Simulated Annealing Structures of CAPPD* with the Lowest Overall Energies

<table>
<thead>
<tr>
<th>Ramachandran Plot Statistics (%)</th>
<th>Average r.m.s. Deviations from Idealized Covalent Geometry (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>residues in most favored regions</td>
<td>95.1</td>
</tr>
<tr>
<td>residues in additionally allowed regions</td>
<td>3.9</td>
</tr>
<tr>
<td>residues in generously allowed regions</td>
<td>1.00</td>
</tr>
<tr>
<td>residues in nonallowed regions</td>
<td>0.00</td>
</tr>
<tr>
<td>backbone residues 460–569</td>
<td>0.95 ± 0.28</td>
</tr>
<tr>
<td>heavy atom residues 460–569</td>
<td>1.67 ± 0.34</td>
</tr>
<tr>
<td>backbone secondary structure</td>
<td>0.40 ± 0.14</td>
</tr>
<tr>
<td>heavy atom secondary structure</td>
<td>1.24 ± 0.3</td>
</tr>
</tbody>
</table>

a None of the structures have distance violations of >0.1 Å or dihedral angle violations of >0.5°. b Determined by PROCHECK (24). c Four α-helices (residues 461–482, 487–518, 526–549, and 552–566).

Attempts to measure residual dipolar couplings to better define the relative orientation of the bottom region of the domain were hampered by the conditions used for structure determination.
APP has been the subject of particularly intense studies because of its involvement in Alzheimer’s disease (AD), one of the most common forms of progressive cognitive failure in humans. Elucidating the function of APP imposes a great scientific challenge. Despite extensive experimentation that spans techniques ranging from exhaustive biochemistry to genetic studies, the normal physiological function of APP is still elusive.

...
tional studies will be required to determine whether CAPPD* is directly involved in F-spondin binding and/or APP processing. The atomic structure of CAPPD* described here provides an initial framework for understanding the function of this domain. It will also be important for uncovering the molecular mechanisms of ligand binding to this region and for the analysis of the potential effects of such interactions on APP processing. The fact that the functionally redundant APP homologues, APLP1 and APLP2, also bind F-spondin (8) suggests that it may be feasible to search for chemical agents that interact with the CAPPD and differentially affect APP/APLPs cleavage. This could potentially lead to a new approach for developing inhibitors of Ab production without interfering with the essential biological function(s) of these proteins. The atomic structure of the CAPPD* may prove to be instrumental in discovering such agents.

ACKNOWLEDGMENT

Mass spectrometry analysis of trypsin-resistant fragments was performed at the Protein Chemistry Technology Center at the University of Texas Southwestern Medical Center. N-Terminal sequencing of the fragments was performed at The Protein Chemistry Laboratory at Texas A&M University (College Station, TX).

SUPPORTING INFORMATION AVAILABLE

A full-length sequence alignment of multiple members of the APP/APLP family (Supporting Information Figure 1) and a superposition of 1H—15N HSQC spectra of CAPPD* in the presence or absence of 300 mM GdnHCl and higher NaCl concentrations (Supporting Information Figure 2). This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES