Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Cloning, expression, crystallization and initial crystallographic analysis of the C-terminal domain of the amyloid precursor protein APP

Cora Keil, Robert Huber, Wolfram Bode* and Manuel E. Than*

Max-Planck-Institut für Biochemie, Abteilung Strukturforschung, Am Klopferspitz 18A, 82152 Martinsried, Germany

Correspondence e-mail: bode@biochem.mpg.de, than@biochem.mpg.de

Received 24 May 2004
Accepted 23 June 2004

Alzheimer's disease is associated with typical brain deposits (senile plaques) consisting mainly of neurotoxic amyloid β-peptides. These are proteolytically derived from the large type I transmembrane protein amyloid precursor protein (APP), which is possibly involved in signal transduction. The large C-terminal domain CAPPD of the human APP ectodomain has been cloned, expressed in large amounts in *Escherichia coli* and purified to homogeneity. Well diffracting tetragonal crystals have been obtained and native data have been collected to 2.1 Å. Initial experimental phases from a three-wavelength MAD experiment using $(NH_4)_2OsCl_6$-derivatized crystals are of good quality and show mostly α-helical conformations.

1. Introduction

Alzheimer's disease (AD) is the most commonly occurring dementia worldwide, especially in the elderly population. Its key pathological hallmark is the deposition of senile plaques of neurotoxic amyloid β-peptides (Aβ), which are believed to be the causative agent of AD. These peptides are proteolytically generated from the β-amyloid precursor protein (APP) *via* sequential cleavages by the aspartyl proteinase β-secretase and the large transmembrane proteinase complex γ-secretase. While the pathological steps leading to the formation of Aβ and its subsequent deposition are relatively well understood, only very limited information is available regarding the physiological function of APP (for reviews, see Koo, 2002; Selkoe, 2002; Haass, 2004).

Several splice variants have been reported for APP. In neuronal cells, APP exists predominantly as a 695-residue type I transmembrane protein (Haass & Selkoe, 1993; Selkoe, 2002) consisting of a large ectodomain, a transmembrane domain and a short cytoplasmic tail (AICD). Together with the adaptor protein Fe65 and the histone acetyltransferase Tip60, this latter tail forms a multimeric complex, which is possibly involved in the modulation of transcription and/or gene expression (Cao & Südhof, 2001, 2004; Baek *et al.*, 2002). Based on these and other interactions between AICD and various binding partners (reviewed in Van Gassen *et al.*, 2000; Koo, 2002) as well as the homology between the proteolytic processing of APP and Notch (reviewed in, Steiner & Haass, 2001; Selkoe & Kopan, 2003), the physiological function of APP seems to be in signal transduction, possibly related to Ca^{2+} homeostasis (Leissring *et al.*, 2002).

By analogy to Notch, it is believed that ligand binding to the large ectodomain is followed by signal transduction to intracellular effector proteins *via* the proteolytic release of AICD. In contrast to the small intracellular domain, the large ectodomain of APP exhibits a clearly defined three-dimensional structure composed of several subdomains that could represent independently folded entities. These include (in sequential order, starting from the N-terminus) (i) a signal peptide, required for correct processing and membrane insertion, (ii) the cysteine-rich domain (CRD) containing an N-terminal growth-factor-like folded heparin-binding site (Rossjohn *et al.*, 1999) and a copper-binding domain (Barnham *et al.*, 2003), (iii) a so-called acidic region, (iv) a large C-terminal or central domain (CAPPD) and (v) a potentially unstructured linker region of roughly 100 amino-acid residues directly preceding the transmembrane helix. This linker also contains the β-secretase cleavage site. In addition, further heparin- and zinc-binding regions have been described (Multhaup *et al.*, 1994; Bush *et al.*, 1994; Beher *et al.*, 1996). Atomic structures of APP are known for the N-terminal 96-residue fragment (Rossjohn *et al.*, 1999), for the Kunitz-type proteinase inhibitor (KPI) contained only in the 751- and 770-amino-acid non-neuronal APP isoforms (Hynes *et al.*, 1990; Scheidig *et al.*, 1997) and for a 66-amino-acid fragment containing the copper-binding motif of the CRD (Barnham *et al.*, 2003). However, none of these partial structures displays potential ligand-binding sites, nor do they provide

structural data for the large C-terminally located CAPPD, which contains the binding site of the recently identified APP ligand F-spondin (Ho & Südhof, 2004), the fibroblast-growth-promoting Kang sequence RERMS (Ninomiya *et al.*, 1993) and the collagen-binding site (Beher *et al.*, 1996).

Here, we report the cloning, expression, purification and crystallization of the large C-terminal domain of the APP ectodomain comprising residues Ser295–Asp500 as well as the initial crystallographic analysis of the resulting tetragonal crystals.

2. Materials and methods

2.1. Cloning, expression and purification

A recombinant form of the large C-terminal part of the extracellular domain of APP comprising residues Ser295–Asp500 (neuronal APP_{695} numbering) has been prepared. Briefly, the DNA of interest was amplified by PCR from cDNA coding for the entire APP molecule using the primers 5′-AAA AAA AAA AAA CAT ATG AGT ACC CCT GAT GCC GTT G-3′ and 5′-AAA AAA AAA AAG CTT ACG ACC TTC GAT GTC ATC TGA ATA GTT TTG-3′ and subcloned into pET22b (Novagene, Darmstadt, Germany), resulting in plasmid pCK3, which also codes for an N-terminal addition of methionine and the C-terminal extension by the His-tag sequence Ile-Glu-Gly-Arg-Lys-Leu-Ala-Ala-Ala-Leu-Glu-His-His-His-His-His-His. For expression, pCK3 was transfected into various *Escherichia coli* cell lines, including BL21(DE3) (Novagen), BL21-CodonPlus(DE3)-RIL (Stratagene, La Jolla, CA, USA) and Rosetta(DE3) (Novagen). Cells were grown in Luria–Bertani medium containing the respective antibiotics, induced at OD_{600} = 0.8 with 1 m*M* isopropyl *β*-thiogalactopyranoside (IPTG), harvested by centrifugation 4 h after induction and lysed by sonification after resuspension in buffer *A* (100 m*M* Tris–HCl, 500 m*M* NaCl pH 6.8).

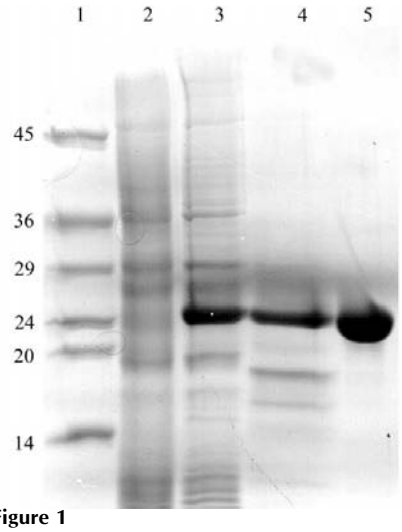

Figure 1
Expression and purification of the C-terminal (Ser295–Asp500) fragment of the APP ectodomain. The following samples were applied to a 12% SDS–PAGE and stained with Coomassie blue: lane 1, molecular-weight markers; lanes 2 and 3, whole cell homogenate from non-induced and IPTG-induced cultures, respectively; lane 4, protein after Ni–NTA purification; lane 5, protein after gel filtration. The IPTG-inducible band runs at the expected molecular weight of 25 kDa and can easily be purified to homogeneity. The molecular weights of the markers are indicated in kDa on the left.

The cell lysate containing the His-tagged protein was cleared by centrifugation and applied onto an Ni–NTA column (Qiagen, Hilden, Germany) pre-equilibrated with buffer *A*. The loaded column was extensively washed with buffer *B* (buffer *A* plus 50 m*M* imidazole) and the protein was eluted with buffer *C* (buffer *A* plus 250 m*M* imidazole). The protein was concentrated by ultrafiltration and further purified by gel filtration using a Superdex 75 HR 26/60 column (Amersham Biosciences, Uppsala, Sweden) equilibrated in buffer *D* (5 m*M* Tris–HCl, 150 m*M* NaCl pH 8). Protein-containing fractions were combined and concentrated to 10 mg ml^{-1}. N-terminal amino-acid sequencing by Edman degradation and matrix-assisted laser desorption–ionization time-of-flight (MALDI-TOF) mass spectrometry confirmed the size and integrity of the desired protein construct. Protein concentrations were determined using a BioRad Protein assay (Hercules, CA, USA).

Selenomethionine (SeMet) derivatized protein was prepared in an analogous way, except that methionine-auxotrophic B834(DE3) cells (Novagene) carrying the pRIL plasmid (Stratgene) were grown in new minimal media (Budisa *et al.*, 1995) complemented with SeMet as described in Budisa *et al.* (1997).

2.2. Crystallization, data collection and data processing

Crystallization was performed at 293 K in sitting drops by the vapour-diffusion method using 24-well Cryschem plates (Charles Supper Company, MA, USA) and various commercially available and in-house crystallization solutions. 1.5 µl protein solution (10 mg ml^{-1} in buffer *D*) and 1.5 µl precipitant were mixed and equilibrated against 300 µl of reservoir solution.

Initial diffraction images were collected using a rotating-anode X-ray generator (Rigaku, Japan) equipped with a Cu *Kα* source and a MAR image-plate detector (MAR Research, Norderstedt, Germany) and were processed and scaled with *MOSFLM* (Leslie, 1991) and *SCALA* (Collaborative Computational Project, Number 4, 1994).

SeMet-containing crystals were prepared by analogy to the native crystals. All further experiments were performed with the better diffracting SeMet crystals, which proved to be isomorphous to the native crystals. For derivatization, crystals were soaked for 12 h in artificial mother liquor complemented with 5 m*M* $(NH_4)_2OsCl_6$. The crystals were quickly transferred into fresh crystallization buffer supplemented with 30% glycerol as cryoprotectant and were flash-frozen in a Cryostream (Oxford Cryosystems, Oxford, UK). Three complete data sets were recorded at the wiggler beamline BW6 at DORIS, Deutsches Elektronen Synchrotron (DESY, Hamburg, Germany) from $OsCl_6^{2-}$-derivatized crystals at wavelengths of 1.13985 Å (peak), 1.14040 Å (edge) and 1.137 Å (remote). For these MAD data sets, two angular segments of 70° in inverse-beam geometry were collected in frames of 1.0°. For the high-resolution data, two data sets were measured at different exposure times over an angular range of 65° using a wavelength of 1.05 Å and frames of 0.2 and 1.0°. All synchrotron data were collected at 100 K using a MAR CCD (MAR Research) detector.

The data were processed and scaled with the *HKL* package (Otwinowski & Minor, 1997). Self-rotation Patterson searches were calculated with *GLRF* (Tong & Rossmann, 1990). One highly occupied Os site was determined using real-space Patterson search techniques and anomalous difference data of the peak data set in *RSPS* (Knight, 1989). This site was refined and experimental phases were calculated with *MLPHARE* (Otwinowski, 1991) and *SHARP* (de La Fortelle & Bricogne, 1997) using all data from the three-wavelength MAD experiment. The phases determined with *MLPHARE* and *SHARP* were further improved by solvent flattening using *DM* (Collaborative Computational Project, Number 4, 1994) and *SOLOMON* (Abrahams & Leslie, 1996), respectively. Anomalous difference Fourier maps were calculated and analyzed with *FFT* and *PEAKMAX* (Collaborative Computational Project, Number 4, 1994).

3. Results and discussion

The large C-terminal domain of the APP ectodomain (CAPPD), comprising residues

Ser295–Asp500 (APP_{695} numbering), has been cloned into pET22b and its expression has been investigated using various *E. coli* strains. More than 10 mg of soluble protein per litre of bacterial culture could be obtained from BL21-CodonPlus(DE3)-RIL, but not from Rosetta(DE3) or BL21(DE3) cells. The negligible expression level observed with BL21(DE3) could be a consequence of the difference in tRNA-usage frequency between *E. coli* and *Homo sapiens* (Kane, 1995). However, both BL21-CodonPlus(DE3)-RIL and Rosetta(DE3) cells are derived from BL21(DE3) and contain different helper plasmids, which encode the genes of those tRNAs that are frequently used in higher eukaryotes but are rare in *E. coli*. Hence, it seems interesting to note that useful expression levels were achieved with the pRIL but not with the pRARE helper plasmid.

Owing to the hexahistidine tag attached to the C-terminus of the CAPPD construct encoded by pCK3, the purification proved to be straightforward (Fig. 1). After the Ni–NTA chromatography step the protein was already almost pure, with only few distinct impurities of mostly lower molecular weight remaining. These and putative high-molecular-weight aggregates of the desired protein could be removed by gel filtration on Superdex S75, resulting in highly pure and homogenous protein preparations. The elution time clearly corresponded to monomeric CAPPD (data not shown). SeMet-containing protein preparations showed virtually identical results during purification.

The straightforward expression and purification procedure for human CAPPD described above turned out to be advantageous for its crystallization. From extensive initial crystallization screens comprising various commercially available and in-house screens, one successful crystallization condition was established and refined.

Reproducible large tetragonal prisms (Fig. 2*a*) grew within a few days from 100 m*M* phosphate/citrate buffer, 10% 2-propanol, 200 m*M* Li_2SO_4 pH 4.2 from either native or SeMet-containing protein preparations. The largest crystals had the approximate dimensions of 100 × 100 × 400 µm.

Native crystals mounted in capillaries diffracted to a limiting resolution of 3.5 Å using a conventional X-ray source and had unit-cell parameters $a = b = 102.05$, $c = 47.17$ Å, $\alpha = \beta = \gamma = 90^\circ$. This diffraction limit could be extended to 2.1 Å with flash-frozen SeMet-containing crystals using synchrotron radiation (Fig. 2*b*). Initial attempts to solve the phase problem with MAD experiments using underivatized SeMet crystals failed as it was not possible to determine the Se substructure. Because of their better diffraction properties, all further experiments were performed with the SeMet crystals. A complete data set to 2.1 Å resolution was collected and a complete three-wavelength MAD experiment was performed at the Os L_{III} absorption edge of $(NH_4)_2OsCl_6$-derivatized crystals at the wiggler beamline BW6 at DESY (Table 1). As is usual, the unit-cell parameters changed slightly upon heavy-metal binding and flash-freezing.

Table 1
Statistics of data collection and processing.

Values for the last resolution shell are given in parentheses.

	Native	$(NH_4)_2OsCl_6$ derivative		
	High-resolution†	Peak	Edge	Remote
Unit-cell parameters				
a (Å)	100.65	100.81		
c (Å)	46.05	45.94		
Wavelength (Å)	1.05	1.13985	1.14040	1.137
f' (e^-)		−11.1	−30.9	−12.1
f'' (e^-)		27.0	15.9	9.8
Resolution (Å)	20–2.1 (2.13–2.10)	20–2.5 (2.53–2.50)	20–2.5 (2.53–2.50)	20–2.5 (2.53–2.50)
No. reflections				
Measured	77682	85452	87028	87152
Unique	26655	30437‡	30879‡	30963‡
Completeness (%)	97.7 (98.1)	97.4 (99.5)‡	98.5 (99.3)‡	98.7 (99.4)‡
R_{merge}§ (%)	6.1 (45.9)	6.9 (41.4)‡	7.0 (41.0)‡	7.0 (42.0)‡
$\langle I/\sigma(I)\rangle$	12.5 (2.3)	12.1 (2.7)‡	12.2 (2.6)‡	12.1 (2.6)‡

† Statistics are given for merged data from two scans at different exposure times. ‡ For the MAD data sets, Friedel pairs were considered to be independent. § $R_{merge} = \sum_h \sum_i |I_{i,h} - \langle I_h\rangle| / \sum_h \sum_i I_{i,h}$.

Data reduction in *P*4 and *P*422 and calculations of the Matthews coefficient (2.2 $Å^3\ Da^{-1}$, corresponding to a solvent content of 44%; Matthews, 1968) indicated that two CAPPD molecules probably represent the asymmetric unit of the $P4_x$ crystals. In good agreement, Patterson self-rotation searches (data not shown) indicated, in addition to the principal crystallographic fourfold symmetry axis, a local twofold axis in the *xy* plane, deviating by 22° from the *x* axis. Because of the relatively short *z* axis, only few 00*l* reflections could be inspected, which did not allow a clear distinction between *P*4, $P4_{1/3}$ or $P4_2$ symmetry. Hence, all further calculations were performed in parallel considering all four possible space groups.

Harker plots of anomalous difference Patterson maps at $w = 0$, $w = 1/4$ and $w = 1/2$ calculated from the Os-peak data set only indicated very strong peaks on the $w = 1/4$ Harker section, consistent with $P4_{1/3}$. As expected, real-space Patterson searches indicated a clear solution in $P4_1$ (but not in *P*4 or $P4_2$), corresponding to one Os site. Refinement of this site and phase calculations in $P4_1$ resulted in an overall figure of

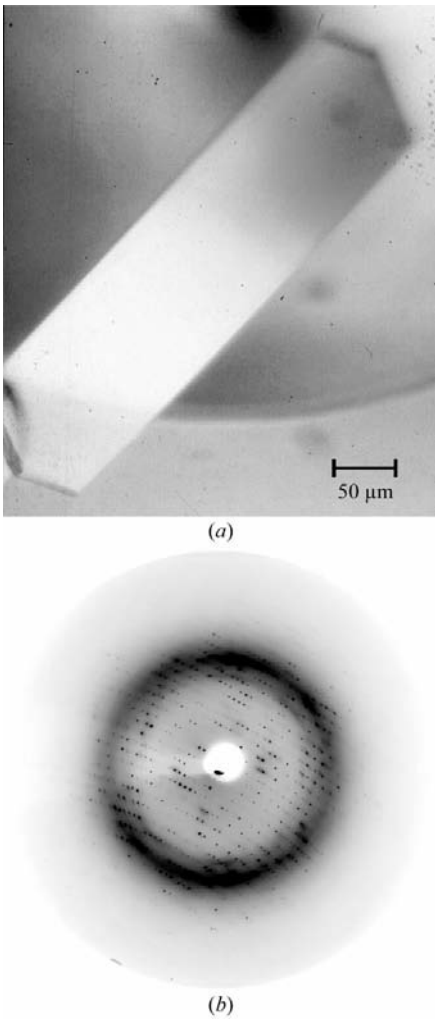

Figure 2
Crystal (*a*) and diffraction image (*b*) of the C-terminal (Ser295–Asp500) fragment of the APP ectodomain. Crystals were typically of tetragonal shape and had maximal dimensions of 100 × 100 × 400 µm. The diffraction image belongs to the high-resolution data set; the edge of the CCD detector area corresponds to 2.0 Å resolution.

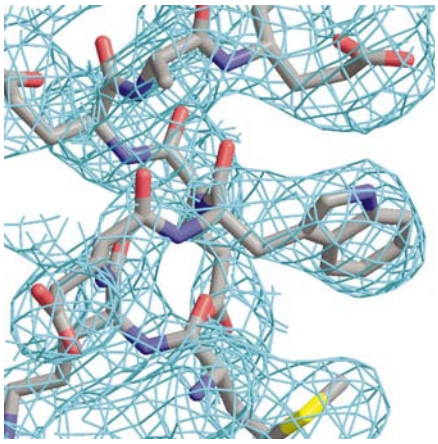

Figure 3
Experimental electron-density map after phasing in *MLPHARE*. A helical region surrounding Trp338 (APP_{695} numbering) is shown at a contour level of 1σ. This figure was prepared using *MOLSCRIPT* (Kraulis, 1991), *BOBSCRIPT* (Esnouf, 1999) and *RASTER3D* (Merritt & Bacon, 1997).

merit (FOM) of 0.43 (20.0–2.70 Å, FOM = 0.53 for the last well phased resolution shell, 3.44–3.99 Å) and a relatively large *B* factor of around 110 Å^2. The resulting electron-density map (Fig. 3) was of good quality and superior to the map calculated in $P4_3$ (using inverted Os coordinates). The $P4_1$ electron-density map displayed mostly helical regions, in agreement with secondary-structure predictions. Solvent flattening did not significantly improve the initial experimental electron-density map, nor did it allow a decision between the two enantiomorphic space groups (overall FOM for *DM* in $P4_1$ was 0.56; that in $P4_3$ was 0.58). Similar results were obtained with phase calculations in *SHARP* and *SOLOMON*. Anomalous difference Fourier maps calculated with the experimental phases using again the Os-peak data set did not reveal further potentially lower occupied Os-binding sites.

In summary, we have established the expression of large amounts of the C-terminal domain CAPPD of human APP in *E. coli* and its purification. Well diffracting crystals could be obtained and complete MAD data from $(NH_4)_2OsCl_6$-derivatized crystals were measured, resulting in a well defined electron-density map. In agreement with secondary-structure predictions, mostly helical regions, some of considerable length, can easily be identified. Further model building and refinement are in progress.

This work was supported by the DFG grant SFB596. We thank G. P. Bourenkov and H. D. Bartunik for support during data collection at BW6 and H. Steiner and C. Haass for providing the β-APP cDNA used in this work and for critical reading of the manuscript.

References

Abrahams, J. P. & Leslie, A. G. W. (1996). *Acta Cryst.* D**52**, 30–42.

Baek, S. H., Ohgi, K. A., Rose, D. W., Koo, E. H., Glass, C. K. & Rosenfeld, M. G. (2002). *Cell*, **110**, 55–67.

Barnham, K. J., McKinstry, W. J., Multhaupt, G., Galatis, D., Morton, C. J., Curtain, C. C., Williamson, N. A., White, A. R., Hinds, M. G., Norton, R. S., Beyreuther, K., Masters, C. L., Parker, M. W. & Cappai, R. (2003). *J. Biol. Chem.* **278**, 17401–17407.

Beher, D., Hesse, L., Masters, C. L. & Multhaup, G. (1996). *J. Biol. Chem.* **271**, 1613–1620.

Budisa, N., Karnbrock, W., Steinbacher, S., Humm, A., Prade, L., Neuefeind, T., Moroder, L. & Huber, R. (1997). *J. Mol. Biol.* **270**, 616–623.

Budisa, N., Steipe, B., Demange, P., Eckerskorn, C., Kellermann, J. & Huber, R. (1995). *Eur. J. Biochem.* **230**, 788–796.

Bush, A. I., Pettingell, W. H. Jr, de Paradis, M., Tanzi, R. E. & Wasco, W. (1994). *J. Biol. Chem.* **269**, 26618–16621.

Cao, X. & Südhof, T. C. (2001). *Science*, **293**, 115–120.

Cao, X. & Südhof, T. C. (2004). *J. Biol. Chem.* **279**, 24601–24611.

Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.

Esnouf, R. M. (1999). *Acta Cryst.* D**55**, 946–950.

Haass, C. (2004). *EMBO J.* **23**, 483–488.

Haass, C. & Selkoe, D. J. (1993). *Cell*, **75**, 1039–1042.

Ho, A. & Südhof, C. (2004). *Proc. Natl Acad. Sci USA*, **101**, 2548–2553.

Hynes, T. R., Randal, M., Kennedy, L. A., Eigenbrot, C. & Kossiakoff, A. A. (1990). *Biochemistry*, **29**, 10018–10022.

Kane, J. F. (1995). *Curr. Opin. Biotechnol.* **6**, 494–500.

Knight, S. (1989). PhD thesis, Swedish University of Agricultural Sciences, Uppsala, Sweden.

Koo, E. H. (2002). *Traffic*, **3**, 763–770.

Kraulis, P. J. (1991). *J. Appl. Cryst.* **24**, 946–950.

La Fortelle, E. de & Bricogne, G. (1997). *Methods Enzymol.* **276**, 472–494.

Leissring, M. A., Murphy, M. P., Mead, T. R., Akbari, Y., Sugarman, M. C., Jannatipour, M., Anliker, B., Müller, U., Saftig, P., De Strooper, B., Wolfe, M., Golde, T. E. & LaFerla, F. M. (2002). *Proc. Natl Acad. Sci. USA*, **99**, 4697–4702.

Leslie, A. G. W. (1991). *Proceedings of the CCP4 Study Weekend. Isomorphous Replacement and Anomalous Scattering*, edited by W. Wolf, P. R. Evans & A. G. W. Leslie, pp. 9–22. Warrington: Daresbury Laboratory.

Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.

Merritt, E. A. & Bacon, D. J. (1997). *Methods Enzymol.* **277**, 505–524.

Multhaup, G., Bush, A. I., Pollwein, P. & Masters, C. L. (1994). *FEBS Lett.* **355**, 151–154.

Ninomiya, H., Roch, J. M., Sundsmo, M. P., Otero, D. A., Saitoh, T. (1993). *J. Cell. Biol.* **121**, 879–886.

Otwinowski, Z. (1991). *Proceedings of the CCP4 Study Weekend. Isomorphous Replacement and Anomalous Scattering*, edited by W. Wolf, P. R. Evans & A. G. W. Leslie, pp. 80–86. Warrington: Daresbury Laboratory.

Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.

Rossjohn, J., Cappai, R., Feil, S. C., Henry, A., McKinstry, W. J., Galatis, D., Hesse, L., Multhaup, G., Beyreuther, K., Masters, C. L. & Parker, M. W. (1999). *Nature Struct. Biol.* **6**, 327–331.

Scheidig, A. J., Hynes, T. R., Pelletier, L. A., Wells, J. A. & Kossiakoff, A. A. (1997). *Protein Sci.* **6**, 1806–1824.

Selkoe, D. J. (2002). *J. Clin. Invest.* **110**, 1375–1381.

Selkoe, D. J. & Kopan, R. (2003). *Annu. Rev. Neurosci.* **26**, 565–597.

Steiner, H. & Haass, C. (2001). *J. Mol. Neurosci.* **17**, 193–198.

Tong, T. & Rossmann, M. G. (1990). *Acta Cryst.* A**46**, 783–792.

Van Gassen, G., Annaert, W. & Van Broeckhoven, C. (2000). *Neurobiol. Dis.* **7**, 135–151.