The molecular basis for apolipoprotein E4 as the major risk factor for late onset Alzheimer's disease


This version is available from Sussex Research Online: http://sro.sussex.ac.uk/id/eprint/83616/

This document is made available in accordance with publisher policies and may differ from the published version or from the version of record. If you wish to cite this item you are advised to consult the publisher's version. Please see the URL above for details on accessing the published version.

Copyright and reuse:
Sussex Research Online is a digital repository of the research output of the University.

Copyright and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable, the material made available in SRO has been checked for eligibility before being made available.

Copies of full text items generally can be reproduced, displayed or performed and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.
Abstract: Apolipoprotein E4 (ApoE4) is one of three (E2, E3 and E4) human isoforms of an alpha-helical, 299-amino acid protein. Homozygosity for the ε4 allele is the major risk factor for developing late onset Alzheimer's disease (AD). ApoE2, ApoE3 and ApoE4 differ at amino acid positions 112 and 158 and these sequence variations may confer conformational differences that underlie their participation in the risk of developing AD. Here, we compared the shape, oligomerisation state, conformation and stability of ApoE isoforms using a range of complementary biophysical methods including small angle X-ray scattering, analytical ultracentrifugation, circular dichroism, X-ray fibre diffraction and transmission electron microscopy. We provide an in-depth and definitive study demonstrating that all three proteins are similar in stability and conformation. However, we show that ApoE4 has a propensity to polymerise to form wavy filaments which do not share the characteristics of cross-beta amyloid fibrils. Moreover, we provide evidence for the inhibition of ApoE4 fibril formation by ApoE3. This study shows that recombinant ApoE isoforms show no significant differences at the structural or conformational level. However, self-assembly of the ApoE4 isoform may play a role in pathogenesis and these results open opportunities for uncovering new triggers for AD onset.
Dear Sheena

Thank you for sending the responses for our submitted manuscript. We have carefully considered all the comments and improved the manuscript following these recommendations. Below we outline our response to each comment.

Reviewer #1: We thank the reviewer for their positive remarks.

Authors need to add some discussions on the limitations of this work. Specifically, apoE proteins in the physiological context are glycosylated and can be lipidated, but these bacterial produced apoE proteins have none of these. Thus, while this work has provided insights on biophysical properties of recombinant apoE isoforms as research tools, these may not be directly assumed to be the case in human physiological or diseased conditions."

Response: We agree with the reviewer on this point and we have added further discussion to explain our results in the context of physiological protein (glycosylated and lipidated). We have also ensured that the use of recombinant protein in aqueous buffer has been made clear. The behaviour of lipidated protein will be important for further studies in the future.

Page 10. In vivo, ApoE is glycosylated and also lipidated [45]. While we have explored the structural conformations adopted by recombinant non-modified ApoE, previous studies have shown that the recombinant ApoE produced in E.coli adopts a very similar conformation and folding to protein produced in adenovirus [46]. Further studies will be necessary to explore the behaviour of lipidated proteins.

Reviewer #2: We thank the reviewer for their supportive comments.

I have a number of comments after reading the manuscript and some criticism of the folding data fitting.

1. It was not clear to me initially why the authors were trying to measure the hydrodynamic radius and Rg of the protein. The results are generally listed without a clear hypothesis. It would be a lot clearer to the reader if the introduction to the manuscript and the different results sections clearly indicated what they are testing. I gathered from the discussion that the previous literature on these proteins has suggested that apoE4 has an expanded radius more characteristic of a molten globule or partially disordered state. The data reported here disagree with this hypothesis. It also states at the beginning of the discussion that previous work has measured the stability of these isoforms to be substantially different. Data presented here suggests otherwise. The paper would be considerably strengthened by making this clear from the start.

Response: We thank the reviewer for their comments and we have now added further background in the introduction and results section to explain our aims. In the context of the literature, our work provides an in depth characterisation showing that the recombinant ApoE isoforms differ very little from one another at the level of multimerisation.

Page 5. In contrast to previous studies [25, 26], we show that the three recombinant isoforms share very similar quaternary, tertiary and secondary structures and thermal and chemical stability.
We have developed a method to produce and maintain full length ApoE proteins recombinantly (see methods and [30]) and it was therefore necessary to first fully characterise these proteins.

2. With regards to the data fitting of the GuHCl unfolding data: while there is nothing wrong with the 3-state model proposed and the shape of the unfolding curves of each protein does indeed suggest that there are at least 3 states of the protein, the different free energies for unfolding reported in table 4b are very unlikely to be accurate. The unfolding curves for the three proteins are essentially superimposable so obtaining such hugely different values for the free energies stems from an inaccurate fitting of the transitions. The authors do not report m-values for unfolding and these would provide the reader with an estimate for the accuracy of the fitting. Proteins with similar sizes and shapes will give a similar sum of the m-values because the m value is representative of the total buried hydrocarbon in a folded protein. If the values obtained differ between similar proteins, then it suggests that the model that is fitted is incorrect rather than that the proteins are actually different in terms of amount of buried hydrocarbon. By this I mean that the 3 state model may not be the correct model: the presence of further states of the protein would make the observed transitions look artificially broad, lowering the fitted m-values. Given that the authors have established that all three proteins have the same size and shape, they can happily assume the same m-values for their proteins. I suggest the authors attempt fitting using the same m-values for all 3 apoE isoforms, or if this does not work, they should admit that the data do not fit to a model that assumes 2 unfolding transitions. They would still be justified in reporting the midpoint concentrations of GdnHCl for each apparent transition. The small differences in the values of these observed midpoints for each isoform are a better reflection of any potential differences in stability. As concluded in the text, the stability difference between the isoforms is indeed marginal and does not seem to explain the fibrillisation propensity of the proteins.

Response: We thank the reviewer for their helpful suggestions. We find that the data does not fit to a 2 state model and agree that the data in the original table was extensive and parameters likely co-vary resulting in inaccurate values. Therefore, we have now replaced this table with a table showing transition midpoints as they are very well defined by the data and altered the results text to refer to this new table (now named table 4c).

3. The different rates of assembly could perhaps be quantified with a t1/2 (rough estimate may be possible using current data?) so that any difference in stability and therefore potentially the unfolding rate could be compared to these values. It is difficult to make a clear link in the absence of a known concentration dependence of assembly.

Response: The data provided shows very obvious and clear differences in the assembly propensity of the three isoforms and we do not feel that calculation of t1/2 would provide further information. Furthermore, electron microscopy supports this view, showing fibrillar assemblies from ApoE4 but not from E3 or E2 during the incubation time and conditions used. Furthermore, our data supports the view that extensive unfolding of the ApoE4 protein is unnecessary for assembly to take place. Therefore, we don’t have a clear link between unfolding (and stability) and assembly in this case.

4. Top of page 11, the authors state that all proteins have the same level of unfolding at 37 degrees. While this is apparent from the CD data, the stability data provides a measure of how many unfolded (U) or partially unfolded (I) molecules are present at any temperature. While the proportion of U or I is of course very small at 37 degrees, it will still be different to the same extent for each isoform, such that a sample of ApoE4 will potentially have 10x more U or I molecules than ApoE2. This may lower the barrier to any assembly state as each molecule of ApoE4 will then be 10x more likely to unfold, even if these events are rare at 37 degrees.

Response: This is an interesting point and thank the reviewer for their insightful comments. We have now calculated the % of protein unfolded at 37°C based on the best fit unfolding curves to our data and included this information in a new table 4b. This shows that there are minor differences
between the amounts of ApoE3, E4 and E2 unfolded at 37°C but these are within experimental error so we do not see any significant difference at physiological temperature.

5. To follow on from point 4., the small amount of GuHCl required to induce a first unfolding transition suggests that these proteins are only marginally stable at the temperature of the experiment. With regards to this temperature, it is unclear what conditions the protein is equilibrated in as the GuHCl titration samples are incubated overnight at 4 degrees then measured in a fluorimeter at 20 degrees. Depending on the folding and unfolding rates of the protein, this may or may not be sufficient to re-equilibrate the samples to the new temperature. Do the titrations relate to the stability of the proteins at 4 degrees or at 20 degrees? Given that assembly is then measured at 37 degrees, it may be useful to understand how the stability of the individual domains changes at this temperature, something which would require titration with GuHCl at 37 degrees if at all possible.

Response. We are grateful for these suggestions. However, our self-assembling protein presents some difficulties in following chemically induced unfolding. The methods explain that the proteins are incubated at 4°C overnight and is then incubated for 5 mins at 20°C during the five scans collected which are then averaged. We do not observe any difference between these graphical outputs. The volume is small and therefore would expect temperature equilibration to take places rapidly. Unfortunately, ApoE proteins are prone to fragmentation and degradation, so it is important to maintain them at a low temperature over time (hence the choice of 4°C overnight). Furthermore, we have observed that at 37°C, the three proteins do not appear to differ in their degree of unfolding and that overtime, ApoE4 assembles. Here we are interested to examine and compare the unfolding of the three isoforms and therefore, we feel that the experimental design is suitable for this work to avoid any deterioration or aggregation of the samples.

Thank you for your consideration.

With best wishes,

Louise
Dear Sheena

Thank you for sending the responses for our submitted manuscript. We have carefully considered all the comments and improved the manuscript following these recommendations. Below we outline our response to each comment.

Reviewer #1: We thank the reviewer for their positive remarks.

Authors need to add some discussions on the limitations of this work. Specifically, apoE proteins in the physiological context are glycosylated and can be lipidated, but these bacterial produced apoE proteins have none of these. Thus, while this work has provided insights on biophysical properties of recombinant apoE isoforms as research tools, these may not be directly assumed to be the case in human physiological or diseased conditions.

Response: We agree with the reviewer on this point and we have added further discussion to explain our results in the context of physiological protein (glycosylated and lipidated). We have also ensured that the use of recombinant protein in aqueous buffer has been made clear. The behaviour of lipidated protein will be important for further studies in the future.

Page 10. In vivo, ApoE is glycosylated and also lipidated [45]. While we have explored the structural conformations adopted by recombinant non-modified ApoE, previous studies have shown that the recombinant ApoE produced in E.coli adopts a very similar conformation and folding to protein produced in adenovirus [46]. Further studies will be necessary to explore the behaviour of lipidated proteins.

Reviewer #2: We thank the reviewer for their supportive comments.

I have a number of comments after reading the manuscript and some criticism of the folding data fitting.

1. It was not clear to me initially why the authors were trying to measure the hydrodynamic radius and Rg of the protein. The results are generally listed without a clear hypothesis. It would be a lot clearer to the reader if the introduction to the manuscript and the different results sections clearly indicated what they are testing. I gathered from the discussion that the previous literature on these proteins has suggested that apoE4 has an expanded radius more characteristic of a molten globule or partially disordered state. The data reported here disagree with this hypothesis. It also states at the beginning of the discussion that previous work has measured the stability of these isoforms to be substantially different. Data presented here suggests otherwise. The paper would be considerably strengthened by making this clear from the start.

Response: We thank the reviewer for their comments and we have now added further background in the introduction and results section to explain our aims. In the context of the literature, our work provides an in depth characterisation showing that the recombinant ApoE isoforms differ very little from one another at the level of multimerisation.

Page 5. In contrast to previous studies [25, 26], we show that the three recombinant isoforms share very similar quaternary, tertiary and secondary structures and thermal and chemical stability and

We have developed a method to produce and maintain full length ApoE proteins recombinantly (see methods and [30]) and it was therefore necessary to first fully characterise these proteins.

2. With regards to the data fitting of the GuHCl unfolding data: while there is nothing wrong with the 3-state model proposed and the shape of the unfolding curves of each protein does indeed suggest that there are at least 3 states of the protein, the different free energies for unfolding reported in table 4b are very unlikely to be accurate. The unfolding curves for the three proteins are essentially superimposable so obtaining such hugely different values for the free energies stems from an inaccurate fitting of the transitions. The authors do not report m-values for unfolding and these would provide the reader with an estimate for the accuracy of the fitting. Proteins with similar sizes and shapes will give a similar sum of the m-values because the m value is representative of the total buried hydrocarbon in a folded protein. If the values obtained differ between similar proteins, then it suggests that the model that is fitted is incorrect rather than that the proteins are actually different in terms of amount of buried hydrocarbon. By this I mean that the 3 state model may not be the correct model: the presence of further states of the protein would make the observed transitions look artificially broad, lowering the fitted m-values. Given that the authors have established that all three proteins have the same size and shape, they can happily assume the same m-values for their proteins. I suggest the authors attempt fitting using the same m-values for all 3 apoE isoforms, or if this does not work, they should admit that the data do not fit to a model that assumes 2 unfolding transitions. They would still be justified in reporting the midpoint concentrations of GdnHCl for each apparent transition. The small differences in the values of these observed midpoints for each isoform are a better reflection of the protein would make the observed transition.
of any potential differences in stability. As concluded in the text, the stability difference between the isoforms is indeed marginal and does not seem to explain the fibrillisation propensity of the proteins.

Response: We thank the reviewer for their helpful suggestions. We find that the data does not fit to a 2 state model and agree that the data in the original table was extensive and parameters likely co-vary resulting in inaccurate values. Therefore, we have now replaced this table with a table showing transition midpoints as they are very well defined by the data and altered the results text to refer to this new table (now named table 4c).

3. The different rates of assembly could perhaps be quantified with a t1/2 (rough estimate may be possible using current data?) so that any difference in stability and therefore potentially the unfolding rate could be compared to these values. It is difficult to make a clear link in the absence of a known concentration dependence of assembly.

Response: The data provided shows very obvious and clear differences in the assembly propensity of the three isoforms and we do not feel that calculation of t1/2 would provide further information. Furthermore, electron microscopy supports this view, showing fibrillar assemblies from ApoE4 but not from E3 or E2 during the incubation time and conditions used. Furthermore, our data supports the view that extensive unfolding of the ApoE4 protein is unnecessary for assembly to take place. Therefore, we don’t have a clear link between unfolding (and stability) and assembly in this case.

4. Top of page 11, the authors state that all proteins have the same level of unfolding at 37 degrees. While this is apparent from the CD data, the stability data provides a measure of how many unfolded (U) or partially unfolded (I) molecules are present at any temperature. While the proportion of U or I is of course very small at 37 degrees, it will still be different to the same extent for each isoform, such that a sample of ApoE4 will potentially have 10x more U or I molecules than ApoE2. This may lower the barrier to any assembly state as each molecule of ApoE4 will then be 10x more likely to unfold, even if these events are rare at 37 degrees.

Response: This is an interesting point and thank the reviewer for their insightful comments. We have now calculated the % of protein unfolded at 37°C based on the best fit unfolding curves to our data and included this information in a new table 4b. This shows that there are minor differences between the amounts of ApoE3, E4 and E2 unfolded at 37°C but these are within experimental error so we do not see any significant difference at physiological temperature.

5. To follow on from point 4., the small amount of GuHCl required to induce a first unfolding transition suggests that these proteins are only marginally stable at the temperature of the experiment. With regards to this temperature, it is unclear what conditions the protein is equilibrated in as the GuHCl titration samples are incubated overnight at 4 degrees then measured in a fluorimeter at 20 degrees. Depending on the folding and unfolding rates of the protein, this may or may not be sufficient to re-equilibrate the samples to the new temperature. Do the titrations relate to the stability of the proteins at 4 degrees or at 20 degrees? Given that assembly is then measured at 37 degrees, it may be useful to understand how the stability of the individual domains changes at this temperature, something which would require titration with GuHCl at 37 degrees if at all possible.

Response: We are grateful for these suggestions. However, our self-assembling protein presents some difficulties in following chemically induced unfolding. The methods explain that the proteins are incubated at 4°C overnight and is then incubated for 5 mins at 20°C during the five scans collected which are then averaged. We do not observe any difference between these graphical outputs. The volume is small and therefore would expect temperature equilibration to take places rapidly. Unfortunately, ApoE proteins are prone to fragmentation and degradation, so it is important to maintain them at a low temperature over time (hence the choice of 4°C overnight). Furthermore, we have observed that at 37°C, the three proteins do not appear to differ in their degree of unfolding and that overtime, ApoE4 assembles. Here we are interested to examine and compare the unfolding of the three isoforms and therefore, we feel that the experimental design is suitable for this work to avoid any deterioration or aggregation of the samples.
Highlights

- There are three Apolipoprotein E isoforms and E4E4 increases the risk for Alzheimer’s disease
- ApoE2, E3 and E4 share a quaternary, tertiary and secondary structures, and stability
- ApoE4 forms non-amyloid-like filaments, while ApoE2 and E3 do not.
- Assembly of E4 may play a role increased risk of late-onset Alzheimer’s disease.
The molecular basis for Apolipoprotein E4 as the major risk factor for late onset Alzheimer’s disease

Ana-Caroline Raulin¹*, Lucas Kraft²*, Youssra K. Al-Hilaly³, Wei-Feng Xue³, John E. McGeehan⁴, John R. Atack²a, Louise Serpell¹x

*authors contributed equally to this manuscript

1. Sussex Neuroscience, School of Life Sciences, University of Sussex, Falmer, Brighton, East Sussex, BN1 6NN, UK
2. Sussex Drug Discovery Centre, School of Life Sciences, University of Sussex, Falmer, Brighton, East Sussex, BN1 6NN, UK
3. School of Biosciences, University of Kent, CT2 7NJ
4. School of Biological Sciences, Institute of Biological and Biomedical Sciences, Faculty of Science, University of Portsmouth, Portsmouth, Hampshire, PO1 2DY, UK

a Current address: Medicines Discovery Institute, Cardiff University, Cardiff, CF10 3AT, UK
x corresponding author

Running title: Structural basis for increased AD risk in ApoE4 genotype
Abstract

Apolipoprotein E4 (ApoE4) is one of three (E2, E3 and E4) human isoforms of an α-helical, 299-amino acid protein. Homozygosity for the ε4 allele is the major risk factor for developing late onset Alzheimer’s disease (AD). ApoE2, ApoE3 and ApoE4 differ at amino acid positions 112 and 158 and these sequence variations may confer conformational differences that underlie their participation in the risk of developing AD. Here, we compared the shape, oligomerisation state, conformation and stability of ApoE isoforms using a range of complementary biophysical methods including small angle X-ray scattering, analytical ultracentrifugation, circular dichroism, X-ray fibre diffraction and transmission electron microscopy. We provide an in-depth and definitive study demonstrating that all three proteins are similar in stability and conformation. However, we show that ApoE4 has a propensity to polymerise to form wavy filaments which do not share the characteristics of cross-β amyloid fibrils. Moreover, we provide evidence for the inhibition of ApoE4 fibril formation by ApoE3. This study shows that recombinant ApoE isoforms show no significant differences at the structural or conformational level. However, self-assembly of the ApoE4 isoform may play a role in pathogenesis and these results open opportunities for uncovering new triggers for AD onset.

199 words

Key words: Apolipoprotein E; Alzheimer’s disease, small-angle X-ray scattering, analytical ultracentrifugation, alpha-helix
Background

Alzheimer’s disease (AD) is the most prevalent dementia, and the sporadic, late onset form comprises 95% of all AD cases. Diagnosis is based on the observation of specific brain pathology of extracellular amyloid plaques composed of Amyloid-beta (Aβ) peptide and intracellular neurofibrillary tangles formed by tau [1]. The biggest risk factor for developing AD is age, but the ε4 variant of the apolipoprotein E (APOE) gene is the strongest genetic risk factor for the development of late onset AD. The human APOE gene encodes three major protein isoforms, ApoE2, ApoE3 or ApoE4 that differ from one another at amino acid positions 112 and 158. While E2 contains two cysteine residues, E3 contains a cysteine and an arginine, and E4 contains two arginine residues at both sites respectively [2]. ApoE4 presents a risk in a gene dose-dependent manner, with ε3/ε4 heterozygotes having a risk increased by three times and ε4/ε4 homozygotes having up to 15 times more chance of developing AD [3]. ApoE3 is the most common isoform, associated with a neutral risk. ApoE2 is under-represented in the population but is thought to be associated with a lower propensity for AD [4].

ApoE is a predominantly α-helical protein with 299 amino acids that is mainly produced in the liver and by astrocytes in the brain, as well as neurons particularly under stress conditions [5, 6]. ApoE is a component of lipoprotein particles [7] and associates with cholesterol, triglycerides and phospholipids [8]. It is found in the central nervous system participating in high-density lipoproteins (HDL) and in plasma associated with very low-density lipoproteins (VLDL), chylomicron remnants and HDL [9]. It has been suggested that the ApoE isoforms have different lipid binding preferences and that lipid binding may be mediated by a conformational change of the monomer [10]. The ApoE isoform amino acid substitutions are presumed to affect ApoE structure and function and have been suggested to be responsible for ApoE4 homozygotes increased risk for AD [11]. However, the mechanism(s) whereby ApoE4 confers an increased risk of developing AD, whether by a loss-of-function or a gain-of-toxic-function, remain poorly understood. It has been shown that ApoE forms complexes with Aβ through regions within their heparin-binding sites in vitro [12] and was found to be associated with Aβ senile plaques in human brain [13] as well as with neurofibrillary tangles [14]. Histopathological
examination of post-mortem AD brains found a positive correlation between plaque density and dosage of the ε4 allele [14-16] and positron emission tomography studies matched these findings with ε4 carriers having higher Aβ deposition compared to non-carriers [17, 18]. ApoE4 has therefore been suggested to enhance Aβ’s fibrillization and reduce Aβ clearance which is thought to contribute to AD progression [19]. Especially, aggregated forms of ApoE4 seem to enhance Aβ fibrillization and antibodies that specifically recognize non-lipidated and aggregated ApoE have been shown to reduce Aβ deposition in APPPS1-21/APOE4 mice [20]. Recent work has also shown that human ApoE4 transgenic animals have increased deposition of tau compared to ApoE3 human transgenics [21] indicating a link to accumulation of neurofibrillary tangles. Additionally, human induced pluripotent stem cell derived neurons expressing ApoE4 were shown to have higher levels of tau phosphorylation as well as increased degeneration of GABAergic neurons that could be rescued by gene editing and converting ApoE4 into ApoE3 [22].

The NMR structure of full-length ApoE3 reveals a large, predominantly α-helical globular protein, with a flexible C-terminal region composed of a short helix (270-277) followed by an unstructured tail (residues 277-299). The structures for N-terminal domains for ApoE2 and ApoE3 have been solved by X-ray crystallography [23, 24] and show α-helical bundles which map closely to the N-terminal domain of E3. It has been proposed that the amino acid substitutions lead to lowered thermal and chemical stability of ApoE4 and higher propensity to form stable intermediate unfolding states characteristic of molten globules arising from the formation of a salt-bridge that links the C and N terminal domain in ApoE4 [25, 26]. A potential consequence of this instability would be an enhanced proteolysis of ApoE4 in neurons, generating fragments that may interfere with cytoskeletal components such as tau protein and neurofilaments [27, 2006, 28]. Decreased stability and molten globule formation has additionally been suggested to drive ApoE4 to form toxic, fibril-like oligomers from full length protein, whereas ApoE2 and ApoE3 do not appear to form high molecular weight species [29].
Here we conducted an in-depth comparison of the structure and stability of the three recombinantly expressed and purified ApoE isoforms using state-of-the-art biophysical techniques and explored ApoE4’s intrinsic property to oligomerise and aggregate. In contrast to previous studies [25, 26], we show that the three recombinant isoforms share very similar quaternary, tertiary and secondary structures and thermal and chemical stability. We reveal that recombinant ApoE4 forms fibrillar structures, however the resulting curvilinear fibrils are fragile and do not share the typical characteristics of cross-β amyloid. Furthermore, to investigate the ApoE4 dose-effect, we explored the effect of co-assembly of ApoE3 and E4. The presence of ApoE3 resulted decreased rate of ApoE4 self-assembly and an inhibition of elongation to long filamentous structures.

Results

All three ApoE isoforms show similar size and shape

We have developed a method to produce and maintain full length ApoE proteins recombinantly (see methods and [30]) and it was therefore necessary to first fully characterise these proteins. Initial purification confirmed that ApoE2, E3 and E4 have identical mobilities of 34 kDa by non-reducing SDS PAGE (Figure S1). Gel filtration and SEC coupled multi-angle light scattering (SEC-MALS) showed multimerization of all three proteins consistent with a tetramer in solution with a molecular weight of 139-151 kDa (Figure 1A & B, Table 1). A hydrodynamic radius of 5.8 – 6.5 nm was calculated for each isoform and a frictional ratio $f/f_0$ above 1.7 suggests elongated shape for the homotetramers (Table 1, Figure S2). Analytical ultracentrifugation was used to investigate the size and shape in solution in further detail. As we extensively dialyzed or buffer exchanged ApoE isoforms into phosphate buffer prior to stability and aggregation studies, we were interested if dialysis affected ApoE oligomerisation in solution. We therefore compared sedimentation in the original size exclusion buffer and in phosphate buffer after dialysis. AUC revealed no differences in sedimentation
velocity between the three isoforms in either buffer (Figure S3, Figure 1 C and D, Table 2a) and the major species was characterised with a molecular mass of 130 kDa (Table 2a).

Small angle X-ray scattering (SAXS) was used to explore whether the tetramers of the ApoE isoforms differ in dimensions and shape. All three proteins exhibited identical scattering profiles indicating that they are very similar in shape (Figure 2A, B, C). A radius of Gyration ($R_g$) of 5.6 nm (56 Å) was calculated using the Guinier approximation and comes close to the hydrodynamic radius determined in the other techniques (Table 2b). The Kratky plot indicates that ApoE consist of several domains that are tethered by linkers with extended conformation. This elongated shape is reproduced in the pair distance distribution function and fits to a maximal dimension of approximately 19.5 nm (195 Å) (Figure 2D).

Together AUC, SEC-MALS and SAXS show that the three ApoE isoforms are similar in size, shape and multimerization, and all form tetrameric species in solution.

**ApoE2, E3 and E4 conformationally similar and show only marginal differences in stability**

Circular dichroism (CD) spectroscopy and tryptophan fluorescence were used to probe potential differences in secondary and tertiary structure of the three isoforms. Negligible spectral differences were observed, suggesting no major conformational differences between the three proteins (Figure 3A and B). Analysis of CD spectra using Dichroweb [31 2002, 32, 33] showed that the three isoforms all possessed around 58% α-helical content (Table 3), with no significant differences between them (one-way ANOVA: F(2,9)=4.197, p>0.05).

Folding stability of ApoE proteins was examined using thermal and chemical denaturation and monitored using CD and tryptophan fluorescence. Thermal denaturation monitored using CD at 222nm showed a sigmoidal unfolding curve for all three isoforms. A phenomenological Boltzmann sigmoidal curve was fitted to the data and revealed that ApoE3 and ApoE4 have similar melting temperature of 52.39°C and 51.32°C respectively, compared to the apparent higher melting
temperature for ApoE2 of 60.28°C (Figure 3C(i)). However, despite these slight differences in the apparent mid-point melting temperature, no differences were observed at physiological temperature of 37°C. One-way ANOVA provided evidence of no significant difference in the mean ellipticity at 222 nm between the three isoforms at 37°C (F(2,9)= 0.9426, p=0.4249; Table 4a). The curves showing fraction unfolded protein for each isoform calculated based on the fitted sigmoidal curves allow direct comparison of the estimated temperature unfolding transitions ((Figure 3C(ii)); Figure S4 A), showing similar thermal denaturation profile for all three isoforms. Furthermore, the fraction unfolded at 37°C showed there were no significant differences between the three isoforms at physiological temperature, consistent with the mean ellipticity at 222 nm at the same temperature (Table 4b).

Chemical denaturation in guanidium hydrochloride (GuHCl) was performed in reducing conditions (1 mM DTT) and the wavelength of maximum fluorescence emission intensity $\lambda_{\text{max}}$ was plotted against corresponding concentrations of GuHCl. All three denaturation curves demonstrate multi-phasic behaviour and a two-state unfolding model did not fit the denaturation data. Instead, a three-state unfolding model was fitted to the data (Figure 3D(i)), suggesting the existence of at least three different states of the protein, and the existence of one or more intermediate states. The data is, therefore, consistent with the separate unfolding of two protein domains. From the fitted curves, GuHCl concentrations corresponding to the midpoint of each transition were calculated (Table 4c).

Similar to the thermal denaturation study, fraction unfolded curves were calculated for each isoform in order to directly compare the shape of the chemical denaturation profile (Figure 3D(ii)). The transition from a folded state to an intermediate state occurs at a slightly lower GuHCl concentration for ApoE4, suggesting that the first domain to unfold denatures marginally more easily for ApoE4 than ApoE2 and ApoE3 (0.58 M vs ~0.71 M). ApoE3 and ApoE4 show similar unfolding of the second domain, while ApoE2 is more stable (~2.65 M vs 2.84 M). Overall, the observed increased chemical denaturant concentration and temperature needed for ApoE2 to unfold compared with ApoE3 and E4 infers that ApoE2 has a slight increased resistance to denaturation compared to E3 and E4 although the difference in apparent stability is marginal (Table 4c).
**ApoE4 assembles to form non-amyloid, native-like fibres**

To further investigate the behaviour of the ApoE isoforms, the three proteins were incubated in 20 mM PB at 37°C for 24 h. Native PAGE showed that ApoE2 and E3 ran consistently at a similar mobility at 0 h and after 24 h incubation. However, ApoE4 also formed higher MW species prior to, and following incubation. By 24 h, the majority of the ApoE4 protein formed higher oligomers that did not run through the gel indicating oligomerisation of the ApoE4 protein (Figure 4A). Thioflavin T (ThT) fluorescence assay is frequently used to monitor molecular self-assembly in solution and although it is often utilised to monitor amyloid formation [34], it is not specific for β-sheet structure or for amyloid [35, 36]. ThT was excited at 440 nm and its fluorescence intensity was followed at 483 nm. The change in fluorescence adjusted to baseline displayed a rapid increase in intensity at 483 nm for ApoE4 (Figure 4B(i&ii)), and an extended lag phase for ApoE3 before any changes were detected (Figure B(ii)). There was very little change in fluorescence for ApoE2 up to 24 h and only minor increase up to 60 h (Figure 4B). After 3 days incubation at 37°C, the ThT fluorescence increased further for ApoE4, while no further differences in ThT intensity were observed for ApoE2 (Figure S5). ThT fluorescence data indicate fibrillisation of ApoE4, while only minor changes and no changes were observed for ApoE3 and E2 respectively. TEM was used to observe morphological changes with time. TEM of ApoE2 showed some small amorphous and some round species after 24 h and 3 days incubation respectively (Figure 4C). Small round species were observed at 24 h for ApoE3 that increased in size over time and became slightly elongated after 3 days incubation (Figure 4C). In contrast, ApoE4 showed fibrillar structures with a curvy appearance at 24 h (Figure 4C).

To further explore the assembly of ApoE4, TEM was used to monitor the size and morphology of the assemblies at different time points up to 24 h. The electron micrographs show accumulation of small rounded species after only 1 h incubation. The formation of short, curved filaments is detected by 3 h incubation, which develop into curvy-linear filaments by 6 h and elongate into extended, curvy filaments by 24 h (Figure 5A). Measurement of the average length of the filaments (Figure 5B) confirms that the filaments extend in length while the diameter of the structures remains invariant.
over 24 h, with an average length of 363 nm and width of 30 nm (Figure 5B and C). Close inspection of the filamentous structures shows that they have a granular appearance, and do not share the twisted and smooth appearance generally observed for amyloid fibrils [37-39]. Fibril formation appears to arise from the end-to-end fusion of the small spherical species, making this unit the apparent smallest building block of the structure.

CD and X-ray fibre diffraction experiments were conducted to investigate the nature of ApoE4 fibrils and to investigate whether these assemblies are amyloid-like in structure and undergo the expected conformational change to β-sheet [29]. CD spectra collected for ApoE4 fibrils incubated for 24 h showed a spectrum with minima at 208 nm and 222 nm consistent with a predominantly α-helical secondary structure content (Figure 5D), of comparable shape and intensities to that of non-assembled ApoE4 (Figure 3A). To ensure that any remaining soluble protein does not dominate the spectra, the sample was centrifuged at high speed to separate supernatant and pellet fractions. The CD data for the resolubilised pellet fraction show E4 retains its α-helical structure after incubation at 37°C for 24 h, and that there is almost no soluble protein left given the very small CD signal in the supernatant fraction (Figure 5D). The fibre pellet was sonicated to ensure that protein was resuspended sufficiently and the spectrum continued to demonstrate α-helical content.

To examine the molecular structure of the mature ApoE4 filaments, 100 µM ApoE4 was incubated for 24 h and aligned to form a partially-aligned fibre bundle. The X-ray fibre diffraction pattern showed a sharp meridional diffraction signal at 4.5 Å with a weaker diffraction signal at 4.1 Å. On the equator, a diffuse, strong signal was observed at 10.1 Å and a weaker reflection at approximately 21 Å (Figure 5E). Comparison of the relative intensities shows that the meridional reflections are weaker than those on the equator which is dissimilar to those usually observed for amyloid fibril patterns where signals arise from cross-β structure at 4.7 Å and 10 Å on perpendicular axes [40]. The diffraction data obtained supports the CD data described above showing that the ApoE4 retains an α-helical conformation in the fibres which is similar to the structure in the soluble protein.
ApoE3 and ApoE4 influence the assembly of one another

ApoE4 alone is able to self-assemble to form non-amyloid-like filaments. However, the majority of individuals with an ε4 allele are heterozygous and have an ε3 allele. To investigate the influence of ApoE3 presence on ApoE4 assembly and vice versa, ApoE3 was incubated with ApoE4 over the course of 24 h to investigate whether they could enhance/seed or inhibit assembly. ApoE3 and ApoE4 were each at 12.5 μM to give a combined ApoE concentration of 25 μM. Assembly over time was compared to assembly of ApoE3 or ApoE4 alone at 12.5 μM (not shown) or 25 μM. Fig 6A shows that ApoE3 reduced the rate of assembly of ApoE4, measured by ThT fluorescence; the ThT kinetics of the mixed sample did not resemble that of E3 alone either. TEM micrographs were produced for the mixed sample and showed the presence of small round and small amorphous species instead of the long mature fibrils observed for ApoE4 alone (Figure 6B).

Discussion

The manner in which APOE genotype leads to an increased risk for late onset AD remains unclear. Some studies have proposed that the three ApoE isoforms differ in their structure and consequently function, which may explain ApoE4’s differential effects on AD pathogenesis [3, 41]. Modelling and fluorescence resonance energy transfer studies have been used to provide evidence of a more open conformation for ApoE4 compared to ApoE3 [11, 42, 43] suggesting that the N-terminal region of ApoE4 is less well tethered to the C-terminal domain than for ApoE3. Others have pointed to differences in stability and conformation [29, 44] and have suggested that ApoE4 is less stable and more prone to aggregation to form amyloid-like fibrils [29]. In vivo, ApoE is glycosylated and also lipidated [45]. While we have explored the structural conformations adopted by recombinant non-modified ApoE, previous studies have shown that the recombinant ApoE produced in E.coli adopts a
very similar conformation and folding to protein produced in adenovirus [46]. Further studies will be necessary to explore the behaviour of lipidated proteins.

Here we have investigated the size, oligomerisation state and shape of recombinant ApoE isoforms in aqueous solution. Our results are in agreement with previous observations [47] and confirm that all three isoforms form an elongated tetramer in solution. SAXS data provides additional information regarding the low-resolution structure of ApoE isoforms in solution and validates the elongated shape in solution with ApoE qualitatively resembling multiple domains that are tethered by flexible linkers. Identical X-ray scattering profiles between ApoE isoforms suggest no major structural differences in macromolecular architecture.

CD reveals that all three isoforms share a similar α-helical content at physiological pH, when measured at 21°C. This is in contrast to other studies, which have reported differences in α-helical content between the isoforms at 15° C, with ApoE4 being the least α-helical protein [44]. Here, temperature and chemical denaturation studies revealed very small differences in stability towards unfolding for the three isoforms. Melting curves displayed a sigmoidal shape that is characteristic for two-state unfolding, with \( T_{m}(E2) \geq T_{m}(E3) \approx T_{m}(E4) \), which is in agreement with previously reported results in terms of order of stability[25, 44, 48]. However, it is also important to highlight that all three isoforms have the very similar secondary structure content and fraction unfolded at physiological temperature of 37°C.

Chemical denaturation displayed multi-phasic unfolding curves for all three isoforms, consistent with the independent unfolding of the C-terminal and N-terminal domains as described by Morrow and colleagues [25]. While the curves cannot be used to fully deduce \( \Delta G^\circ_{\text{H}_2\text{O}} \) and the slope m-value for each transition since the three-state model employed is empirical in this case, we interpret slight differences in the apparent denaturant concentration corresponding to the midpoint of each transitions.

On the one hand, the transition from the folded to intermediate state was very similar for ApoE2 and ApoE3, however, it occurred at a lower denaturant concentration for ApoE4. On the other hand,
transitioning from intermediate to unfolded was most similar between ApoE3 and ApoE4, whereas ApoE2 required a slightly higher denaturant concentration. Others have gone further with chemical denaturation studies by looking at the 10 kDa carboxy- and the 22 kDa amino-terminal domains, in conjunction to the corresponding full-length proteins. Data resulted in the attribution of the first transition to the unfolding of the 10 kDa carboxy fragment and the second to that of the 22 kDa amino-terminal domain, whilst reporting the same order of chemical stability [25, 44, 49].

The loss of stability of ApoE4 as measured by spectroscopic methods, albeit under extreme conditions such as high temperature or under denaturing conditions, may be related to its propensity to polymerise into filamentous structures. We showed here that ApoE4 forms filamentous structures after only 24 h incubation at physiological temperature, pH 7.4, while ApoE2 and ApoE3 remained soluble and globular under the same conditions. TEM reveals that the ApoE4 filaments have a polymeric appearance (beads on a string) and there is a clear hierarchical assembly of small spherical species to small, elongated fibrils and finally fibrils, which show identical diameters. CD shows that following incubation, the ApoE4 fibres retain their α-helical secondary structural content. The CD for the whole fraction is almost identical to ApoE4 prior to incubation. To probe whether the α-helical intensity arises from soluble, unassembled protein, we examined the CD from a sedimented sample. Both pellet and supernatant fractions show a spectrum consistent with high α-helical content. Furthermore, the intensity of the spectrum in the supernatant was very low suggesting that the majority of the ApoE4 protein is found within the fibre containing pellet. This result contrasts with a small shift from α-helix to β-sheet previously described by Hatters et al [29]. Thioflavine T fluorescence assay shows increasing intensity at 483 nm with time showing that the fibril formation course can be followed using this chemical rotamer [36, 50]. Although ThT is well characterised as fluorescent probe for amyloid fibril formation, it is a molecular rotor which is not specifically sensitive to β-sheet structures. Amyloid binding studies have revealed that ThT may associate with aromatic groups including tyrosine ladders [51, 52]. Fluorescence intensity can increase in the presence of many polymeric molecules and in viscous solutions [51]. Furthermore, ThT is fluorescent
in the presence of cross-α fibrous molecules revealing that its fluorescence does not rely on β-sheet structure [53].

X-ray fibre diffraction has been classed as a definitive diagnostic for amyloid fibrils, giving the classical amyloid cross-β diffraction pattern composed of a 4.76 Å meridional arising from hydrogen bonding β-strands running perpendicular to the fibre axis and a more diffuse equatorial signal that arises from β-sheet packing that accommodates side chains [54-56]. Here we revealed that the ApoE4 filaments do not give a cross-β pattern, instead diffraction shows a pattern that can be interpreted as α-helical rich polymer, consistent with the TEM and CD data. An α-helix has a repeat of 5.4 Å parallel to the axis of the helix (1.5 Å rise per residue, with 3.6 residues/turn) [57]. However, the structure of the ApoE molecule is not arranged with all helices arranged in a vertical array. 4.5 Å is a repeating unit along an extended polypeptide chain and was previously observed for dried peptide crystals [58], whilst 10.1 Å approximates to the packing of helices in the vertically aligned ApoE structure. Furthermore, amyloid fibrils are characterised by their high stability and resistance to degradation. For example, amyloid formed by the Alzheimer’s related peptide is SDS resistant [59]. Here we have shown that the ApoE4 fibrils are soluble and the ApoE4 species run on a native gel at similar mobility to ApoE3 and ApoE2. Many examples of native like polymers exist including pathologically related serpins such as α-1-antitrypsin [60] and of course, functional polymeric proteins such as microtubules and actin.

The majority of individuals are heterozygous ε3/ε4, and this confers an increased risk for AD development, albeit lower than for ε4/ε4 individuals [61]. Here we investigated the aggregation potential for mixed population of ApoE3 and ApoE4 and revealed that E3 inhibits E4 fibril formation in vitro, leading to a mid-range aggregation kinetic for assembly. Therefore the increased dose of ApoE4 in ε4/ε4 individuals leads to increased assembly compared to half dose of ApoE4 in ε3/ε4 people.
Conclusions

The detailed studies conducted have revealed that recombinant ApoE2, ApoE3 and ApoE4 resemble one another at a quaternary, tertiary and secondary structural level. All three ApoE isoforms form elongated tetramers in solution and each monomer is rich in α-helical conformation. There is no evidence that the aqueous proteins differ at a structural level. Furthermore, chemical and thermal denaturation studies reveal that the three proteins are similarly stable and follow comparable unfolding mechanisms by melting and by chemical denaturation. However, we have shown that ApoE4 has a higher propensity for polymerisation at physiological temperature and pH, to form elongated fibrous structures which retain the native α-helical conformation. The propensity of ApoE4 to self-assemble may play an important role in the mechanism by which it increases susceptibility to develop AD which impacts on ε4/ε4 individuals more severely than those with ε3/ε4.

Materials and Methods

All materials were purchased from Sigma-Aldrich or Fisher Scientific at the highest purity available.

Protein Production

ApoE2, ApoE3 and ApoE4 were cloned, expressed and purified using methods detailed in the supplementary information. Briefly, a codon optimized ApoE4 gene was cloned into a pET17b vector with a six-histidine tag, thioredoxin (TRX) and HRV 3C protease cleavage site upstream. ApoE2 and ApoE3 genes were generated by site directed mutagenesis using QuikChange. Recombinant ApoE proteins were expressed in E. coli Rosetta2(DE3) cells and affinity purified using Talon ® beads followed by HiTrap heparin affinity column and cleaved on column with PreScission protease. The ApoE proteins were eluted using a salt gradient and further purified using a HiLoad Superdex 26/600 200 pg size exclusion column. ApoE containing fractions were concentrated and stored at -80°C.

Gel filtration studies
The high molecular weight protein standard (GE Healthcare, # 28-4038-42) was used to calibrate the HiLoad Superdex 26/600 200 pg column (GE Healthcare). All protein standards and dextran blue were dissolved in size exclusion chromatography (SEC) buffer (20 mM HEPES, 300 mM NaCl, 10% (v/v) glycerol, pH 8.0) to the manufacturers recommended concentrations (0.4 mg/mL Ferritin, 4 mg/mL Aldolase, 3 mg/mL Conalbumin, 3 mg/mL Ovalbumin) and 1.6 mL of the protein calibration mixture applied onto the column at a flow rate of 1.5 mL/min. Stokes radius (size) values were calculated on the basis of ApoE elution volume (equation (1))

\[ K_{av} = \frac{V_e - V_0}{V_t - V_0} \]  

where \( V_e \) = elution volume for the protein, \( V_0 \) = Superdex 200 void volume (112.26 mL), and \( V_t \) = total Superdex 200 bed volume (320 mL). \( K_{av} \) values were calculated for each protein standard and a standard curve generated by plotting the \( \sqrt{-\log(K_{av})} \) as y values versus the Stoke radius of each standard as x values (Ferritin 6.1 nm, Aldolase 4.8 nm, Conalbumin 4.04 nm, Ovalbumin 2.75 nm).

ApoE Stoke radius \( R_s \) was estimated with the following best fit linear equation (2):

\[ y = 0.1058 x + 0.2397 \quad (R^2 = 0.9552) \]  

ApoE diffusion coefficient \( D \) was estimated from its Stoke radius \( R_s \) by following equation (3):

\[ D = \frac{kT}{6\pi\eta R_s} \]  

where \( k \) is the Boltzmann constant (1.38 x 10^-23 m^2 kg s^-2 K^-1), \( T \) is the experimental temperature (277.15 K) and \( \eta \) is the solvent viscosity (1.4181 x 10^-3 Pa s) at temperature \( T \).

The frictional ratio \( f/f_0 = R_s/R_{\text{min}} \) was calculated by assuming the minimal radius (\( R_{\text{min}} \)) of a sphere that could contain the given mass of tetrameric ApoE (136800 Da) where \( R_{\text{min}} \) is defined as

\[ R_{\text{min}} = 0.066 M^{1/3} \quad (\text{for } M \text{ in Daltons, } R_{\text{min}} \text{ in nm}) \]  

Dialysis and buffer exchange

ApoE was either extensively dialysed overnight in 20 mM phosphate buffer (PB) (16 mM Na₂HPO₄, 4 mM NaH₂PO₄), pH 7.4 using Slide-A-Lyzer™ Dialysis Cassettes with a molecular weight cut-off (MWCO) of at 3.5 kDa (Thermo Fisher) or buffer exchanged using disposable Vivaspin® 500
centrifugal concentrators with a MWCO of 3 kDa or Vivaspin® 20 centrifugal concentrators with a MWCO of 3 kDa or 10 kDa (Sartorius).

Gel Electrophoresis

Protein samples for denaturing one dimensional sodium dodecylsulphate (1D-SDS) – gel electrophoresis were prepared in Laemmli sample buffer (Bio-Rad) to a final concentration of 3 µM, omitting the addition of any reducing agent. Samples were loaded into a 4-20% TGX gel (Bio-Rad), and the gel was run at 120 V (constant voltage) for 80 min. Resolved proteins were stained using a Coomassie dye R-250 containing solution (Imperial protein stain, #24615, Thermo Scientific).

Protein samples for native PAGE were diluted to 3 µM in native sample buffer (Bio-Rad) and were directly loaded into a 4-20% TGX gel (Bio-Rad). Gels were run at 120V for 120 min. Resolved proteins were stained using a Coomassie dye R-250 containing solution.

Small angle X-ray scattering and size exclusion chromatography multi angle light scattering (SEC MALS)

Small angle X-ray scattering (SAXS) and multi angle light scattering (MALS) experiments were performed at the B21 beamline (Diamond Light Source, UK).

X-ray scattering was acquired with an X-ray wavelength of 1 Å on a Pilatus 2M detector at a distance of 3.9 m and a camera length of 4.036 m. ApoE isoforms at 10 mg/mL were delivered at 20°C and a flow rate of 0.16 mL/min via an in-line Agilent high performance liquid chromatography (HPLC) with a Shodex Kw-403 column and 20 mM HEPES, 300 mM NaCl, pH 8.0 as running buffer. In total 620 frames were recorded and each frame exposed for 3s. Buffer subtraction and averaging was performed in ScAtter version 3. Data was analysed in ScAtter version 3 and ATSAS 2.8.3 for comparison.

ApoE isoforms were diluted to 5 mg/mL in 20 mM HEPES, 300 mM NaCl, 1 mM TCEP, pH 8.0 for MALS experiments and delivered at RT via an in-line Agilent HPLC with a Shodex Kw-403 column.
Refractive increments (dn/dc) were determined by a Wyatt optilab T-rEX and scattering measured by a Wyatt Dawn Heleo with QELS. Data was analysed using ASTRA version 6.1.7 (Wyatt).

Analytical ultracentrifugation

Sedimentation velocity experiments were carried out in a Beckman model XL-A analytical ultracentrifuge (Beckman Coulter, Fullerton, CA) at 20ºC using an AnTi60 rotor. ApoE isoforms (400 µL) were used at a concentration of 8 µM in 20 mM HEPES, 300 mM NaCl, 10% (v/v) glycerol, pH 8.0 or at a concentration of 8 µM in 20 mM PB, pH 7.4. Experiments were performed at a rotor speed of 40,000 rpm (128,794 x g) and absorbance at 280 nm was recorded every 25 min, at radial intervals of 0.003 cm. A continuous size distribution c(s) from the Lamm equation model in the range of 0.1S-15S was fit to the data using the programme SEDFIT version 15.01b [63]. The confidence interval (F-ratio) was set to 0.95. Buffer viscosity and density was calculated in Sednterp version 20120828 BETA [64 & Pelletier, 1992].

Circular Dichroism (CD) Spectroscopy

ApoE isoforms were diluted to a final concentration of ~25 µM in 20 mM PB, pH 7.4. The exact concentrations were confirmed in triplicate using a NanoDrop ND2000c spectrophotometer (Thermo Fisher) by measuring the absorbance at 280 nm, and applying Beer-Lambert's law using an extinction coefficient of 44,460 L.mol⁻¹.cm⁻¹. The averaged concentrations were recorded and used in calculations when needed.

FarUV CD data were collected using a Jasco-715 CD-spectrometer (Jasco, Goh-Umstadt, Germany) at different time points. Temperature was maintained at 21ºC using a Peltier controlled cell holder. All spectra were collected in 0.01 cm transparent quartz cuvettes (Starna Scientific, Essex, UK) in the range of 280-180 nm with a resolution of 0.1 nm and a band-width of 1 nm. Scanning-speed was set to 50 nm/min, response time to 4 s and sensitivity to standard. Three spectra per sample were acquired and averaged to give a spectrum after subtracting blank spectra of the buffer. For each isoform, spectra were collected on protein from at least three different production batches and an average trace of [Θ] against wavelength was obtained using GraphPad Prism. The mean residue ellipticity ([Θ] in
deg.cm$^2$.dmol$^{-1}$) was calculated from the measured ellipticity $\Theta$ in mdeg at wavelength $\lambda$ using equation (5):

$$[\Theta] = \frac{\Theta}{10. n. C. l} \quad (5)$$

with $n$ the number of amino-acid bonds in the protein, $C$ the concentration of the sample in mol.L$^{-1}$ and a path length of 1 cm.

Secondary structure of ApoE isoforms was calculated by deconvolution of circular dichroism spectra performed with CONTIN/LL [65, 66] at DichroWeb [32] using the reference spectra set 6 [67]. Results represent a mean of values from a minimum of three spectra per isoform. One-way ANOVA was performed to compare the $\alpha$-helical content between the isoforms, using GraphPad Prism.

**Thermal denaturation using CD**

ApoE isoforms were diluted to a final concentration of ~25 μM in 20 mM PB, pH 7.4. CD data were collected using a Jasco-715 CD-spectrometer (Jasco, Goh-Umstadt, Germany). Temperature was increased at a rate of 1°C/min from 20°C to 80°C using a Peltier controlled cell holder. Thermal scans were acquired at a wavelength $\lambda$ of 222 nm every 1.0°C with a sensitivity of 100 mdeg, a response time of 1s and a bandwidth of 1 nm. Thermal scans were acquired for proteins purified from at least three different production batches and an average trace of $[\Theta]$ against temperature was obtained using GraphPad Prism. The unfolding was irreversible. A phenomenological Boltzmann sigmoidal curve with the following equation (6) was fitted to the averaged thermal denaturation curves:

$$Y = \frac{\text{Top} - \text{Bottom}}{1 + e^{-\frac{\text{Tm} - X}{\text{Slope}}}} \quad (6)$$

with Top = maximum $\Theta$ value, Bottom = minimum $\Theta$ value, Tm the temperature corresponding to 50% change in $\alpha$-helical content and Slope = the steepness of the curve. The calculated thermal denaturation curve based on the fitted curves was also transformed into the fraction of the protein unfolded for comparison of protein transformation with the following equation (7):
Fraction unfolded \( = \frac{\theta_N - \theta}{\theta_N - \theta_D} \) (7)

with \( \theta_N \) the ellipticity of the protein at 20°C corresponding to the folded state and \( \theta_D \) the ellipticity of the denatured protein at 80°C.

**Intrinsic fluorescence measurements**

Recombinant ApoE isoforms were diluted to 10 \( \mu \)M in 20 mM PB and incubated overnight at 4°C. Fluorescence scans were acquired using SpectraMax i3 reader (Molecular dimensions). Excitation was set to 295 nm and scans were collected with 1 nm increments between 320 nm and 465 nm. Excitation and emission bandwidths were set to 9 nm and 15 nm respectively. The number of reading per well was set to 6. The photomultiplier tube voltage was set to high. Blank spectra of the buffer were subtracted to protein fluorescence scans. A minimum of 3 different production batches per isoform was used and an average trace of fluorescence intensity against wavelength was obtained using GraphPad Prism.

**Chemical denaturation and data analysis**

Recombinant ApoE isoforms at \(~0.05\) mg/mL in 20 mM PB, 1 mM dithiothreitol (DTT) were incubated overnight at 4°C with increasing concentrations of guanidine-hydrochloride (GuHCl) (0 to 6 M) [25]. Measurements were obtained with a Varian Cary Eclipse spectrophotometer (Varian Ltd., Oxford, UK), quartz cuvette (1 cm path length; Starna, Essex, UK). Temperature was maintained at 20°C using a Varian Cary temperature controller. Tryptophan residues were selectively excited at 292 nm and emission was monitored between 310 nm and 400 nm; excitation and emission slits were set to 10 nm. Five acquisitions per concentration point were acquired at a scan speed of 90 nm/min using response time of 0.05 s. A minimum of 3 different production batches per isoform were used [68]. Wavelength of maximum emission (\( \lambda_{max} \)) per GuHCl point was determined by peak fitting of the averaged emission spectra [44]. Denaturation curves were obtained by plotting the wavelength corresponding to the maximum fluorescence intensity against GuHCl concentration ([GuHCl]) on
GraphPad Prism. The curves show two major transitions and can be described by a three-state unfolding model [69], as described in equation (8) and (9)

$$
\text{Folded} \rightleftharpoons \text{Intermediate} \rightleftharpoons \text{Unfolded} \quad (8)
$$

$$
Y = \frac{Y_F ([\text{GuHCl}]) + Y_U ([\text{GuHCl}]) \cdot K_{\text{app}}}{1 + K_{\text{app}}} \quad (9)
$$

Equation (9) [69] was fitted to the data with Y corresponding to \( \lambda_{\text{max}} \), and \( Y_F \) and \( Y_U \) the signal corresponding to the folded (F) and unfolded (U) protein respectively. \( Y_F \) and \( Y_U \) are assumed to be linearly dependent on the concentration of denaturant, as displayed in equation (10):

$$
Y_X = Y_X^0 + m_x [\text{GuHCl}] \quad (10)
$$

with X corresponding to either the F or U state and \( Y_X^0 \) the \( \lambda_{\text{max}} \) at either 0 M GuHCl or 6 M. In equation (9) \( K_{\text{app}} \) corresponds to the apparent guanidine-dependent equilibrium constant of the denaturation process, defined by equation (11):

$$
K_{\text{app}} = \frac{K_1 K_2 + AK_1}{1 + (1 - A)K_1} \quad (11)
$$

\( K_1, K_2 \) and \( A \) are defined by equation (12-14)

$$
K_1 = e^{-\frac{\Delta G^\circ_1 \text{H}_2\text{O} - m_1 [\text{GuHCl}]}{RT}} \quad (12)
$$

$$
K_2 = e^{-\frac{\Delta G^\circ_2 \text{H}_2\text{O} - m_2 [\text{GuHCl}]}{RT}} \quad (13)
$$

$$
A = \frac{(Y_I - Y_P^0)}{(Y_U^0 - Y_P^0)} \quad (14)
$$

\( \Delta G^\circ_1 \text{H}_2\text{O}, \Delta G^\circ_2 \text{H}_2\text{O}, m_1 \) and \( m_2 \) were the fitted parameters, while \( \Delta G^\circ_1 \text{H}_2\text{O} \) and \( \Delta G^\circ_2 \text{H}_2\text{O} \) are the linearly extrapolated free energy differences at 0 M GuHCl [69] F to I, and I to U, respectively. Subsequently, the concentrations corresponding to 50% unfolding of the protein to an intermediate state ([GuHCl]_{50, I}) and to an unfolded state ([GuHCl]_{50, U}) were calculated by equations (15) and (16) and compared:

$$
[\text{GuHCl}]_{50, I} = \frac{\Delta G^\circ_1 \text{H}_2\text{O}}{m_1} \quad (15)
$$

$$
[\text{GuHCl}]_{50, U} = \frac{\Delta G^\circ_2 \text{H}_2\text{O}}{m_2} \quad (16)
$$
The fraction unfolded, $A$, was also calculated and plotted as a function of [GuHCl] to facilitate comparison of the shape of the unfolding curves between the three isoforms.

**Thioflavin T (ThT) fluorescence assay of ApoE self-assembly**

ApoE isoforms at ~25 μM in 20 mM PB, pH 7.4 were incubated in the presence of Thioflavin-T (ThT) at 37°C in the SpectraMax i3 plate reader. Aqueous ThT stock solution was prepared at a concentration of 3.14 mM, filtered through a 0.2 μm pore size, and used in a 1:2.2 ratio. 96-well plates were sealed with an optically clear polyolefin film to avoid evaporation (StarSeal Advanced Polyolefin Film, Starlab). The number of reading per well was set to 6, PMT voltage was set to high and blank spectra of the buffer were subtracted to protein fluorescence scans. Excitation wavelength was 440 nm and emission at 483 nm was monitored every 15 min, with 3 s low orbital shakes before readings. Fluorescence intensity at a given time point (F) for each isoform was adjusted by subtracting the minimum fluorescence intensity value (F$_{min}$), and not to the initial fluorescence, to account for the increasing temperature effect on ThT (from room temperature to 37°C).

Adjusted fluorescence was plotted against time, and averaged traces for each isoform was obtained on GraphPad Prism. A minimum of 3 different production batches per isoform was used.

**Negative Stain Transmission Electron Microscopy (TEM)**

Morphology of ApoE after 24 h incubation at 37°C was assessed by negative stain TEM. A droplet of sample (4 μL) was placed on 400-mesh carbon-coated grids (Agar Scientific, Essex, UK) and incubated for 1 min. After blotting the excess solution, the grid was washed with 4 μL filtered Milli-Q water and blotted. It was then negatively stained with 4 μL filtered 0.5% uranyl acetate for 40 s and blotted with filter paper. Grids were left to air-dry for at least 5 min before storage. Grids were examined on a Jeol Jem1400-plus transmission electron microscope (Jeol, U.S.A) operated at 80 kV fitted with a Gatan Orius SC100 camera (UK).
Separation of fibril pellet from supernatant

To probe the secondary structure of ApoE4 fibrils, pellets were isolated from supernatant by ultra centrifugation of assembled fibrils at 60,000 rpm in an OptimaTM MAX Ultracentrifuge (Beckman Coulter) using a Beckman TLA120.2 fixed-angle rotor for 45 min at 4°C. Supernatant was removed and replace by fresh 20 mM PB buffer. CD spectra of the whole fraction, supernatant, pellet and sonicated pellet (5 s sonication) were acquired as described for CD. The data was plotted graphically using GraphPad Prism but without conversion to molar ellipticity. Data for whole fraction, pellet and supernatant for the same sample were compared directly to compare secondary structure and protein content for each fraction.

X-Ray fibre diffraction

ApoE4 (100 µM in PB pH 7.4) was incubated for 24 h at 37°C to form fibrils and then the sample was centrifuged for 40 min at 21,100 rpm at 4°C using a benchtop centrifuge (Mikro 22R, Hettich). The resulting pellet was resuspended in 400 µL filtered milliQ water and this was repeated to remove salts that could interfere with the diffraction pattern. Finally, the pellet was resuspended in 30 µL filtered milliQ water. 10 µL of concentrated fibre sample was suspended between two wax-filled capillaries and incubated in a sealed Petri dish at 4°C to form a partially aligned fibre sample [70]. X-ray diffraction images were collected using a Rigaku rotating anode source (CuKα) and a Saturn CCD+ detector. Partially aligned fibres were placed in the X-ray beam and exposed for 30 s or 60 s at specimen to detector distances of 50 mm and 100 mm. Diffraction patterns were converted to TIFF format using imosfIm [71] and analysed using CLEARER [72].
563 **Abbreviations**
564 Alzheimer’s disease  
565 ApolipoproteinE  
566 Transmission electron microscopy  
567 Small angle X-ray scattering  
568 Analytical ultracentrifugation  
569 Low density lipoproteins  
570 Very low-density lipoproteins  
571 Circular dichroism  
572 thioredoxin  
573 size exclusion chromatography  
574 multi angle light scattering  
575 high performance liquid chromatography  
576 phosphate buffer  
577 one dimensional sodium dodecylsulphate  
578 dithiothreitol  
579 guanidine-hydrochloride  
580 Thioflavine T  
581
582
583 **Declarations**
584 **Ethics approval and consent to participate**
585 There are no ethical considerations related to this work
586
587 **Consent for publication**
588 Not applicable
589
590 **Availability of data and materials**
591 All data generated or analysed during this study are included in this published article and its
592 supplementary information files.
593
594 **Competing Interest**
595 The authors declare they have no conflicts of interest
596
597 **Funding**
LK is funded by the Alzheimer’s society Doctoral training centre grant awarded to LCS at University of Sussex. AR is funded by Sussex Neuroscience doctoral training centre supported by the University of Sussex. The LCS is supported by Alzheimer’s society and Alzheimer’s research UK.

**Author contributions**

LK and AR conducted the work and wrote the paper. LK developed the protein purification and collected and analysed SAXS data. AR collected and analysed folding and assembly data and analysed the folding results with WFX. Both AR and LK collected the AUC data. JM supervised the AUC data collection and managed the analysis. YA collected the diffraction data. LCS and JA managed the research. LCS wrote the paper. JA edited the paper.

**Acknowledgements**

The authors gratefully acknowledge Dr Antony Oliver (GDSC) for help with SAXS analysis, Dr Pascale Schellenberger for transmission electron microscopy. The authors thank Dr Laura Itzaki for help with data analysis. We thank Diamond Light Source for access to beamline B21 that contributed to the results presented here. We thank Robert Rambo, Nathan Cowieson and Nikul Khunti for support during beamtime. The authors thank Caroline Lelliot for her contribution to ThT data collection optimisation.

**References**


Figure 1. Characterisation of recombinant ApoE2, E3 and E4.

A. The SEC elution pattern of ApoE isoforms on the Superdex 200 column as a function of absorbance at 280 nm versus the elution volume. All ApoE isoforms have a minor elution peak at 131 mL and a major at 151 mL suggesting the presence of different oligomeric species. A hydrodynamic radius of 5.8 nm, as well as a diffusion coefficient of \(2.47 \times 10^{-7} \text{ cm s}^{-1}\) is calculated by calibration of the Superdex 200 column (inlet) using commercially available protein standards (1 – ferritin, 2 – aldolase, 3 – conalbumin, 4 – ovalbumin). A frictional ratio \(f/f_0 = R_s/R_{\text{min}}\) of 1.71 is calculated which suggest moderate elongation of ApoE [62]. B. SEC MALS plotted with the differential refractive index (RIU, line) are shown as a function of elution time for ApoE2, ApoE3 and ApoE4. Insert shows the calculated molecular weight (Mw) across the peak. All ApoE isoforms have identical elution volumes and a molecular weight of 145 kDa, 139 kDa and 151 kDa is calculated at elution peak for ApoE2, ApoE3 and ApoE4 respectively. C. and D. AUC showing continuous \(c(S)\) size distributions...
in size exclusion buffer (20 mM HEPES, 300 mM NaCl, 10 % (v/v) glycerol, pH 8.0) (C) and in 20 mM PB pH 7.4 (D). No difference between ApoE isoforms is observed in either buffer condition. A major species with a sedimentation coefficient $S$ at 3 and 5 is observed in size exclusion and PB buffer respectively. The difference in sedimentation coefficient between buffers is due to the presence of glycerol in the storage buffer.

**Figure 2. Small angle X-ray scattering.** X-ray scattering curves and the dimensionless Kratky plot (Sc$\tilde{\text{A}}$tter) for A. ApoE2, B. ApoE3 and C. ApoE4 are shown, as well as D. their corresponding pair distance distribution function $P(r)$. A-C. All ApoE isoforms have identical scattering profiles and adopt an extended conformation in solution with some intrinsic level of flexibility as assessed by the dimensionless Kratky plot. D. This extended conformation is seen in the $P(r)$ distribution respectively and a maximal dimension $D_{\text{max}}$ of approximately 19.5 nm (195 Å) is determined for each isoform.
**Figure 3. Conformation and stability of ApoE isoforms**

A. Far UV CD spectra of ApoE isoforms (25 µM in 20 mM PB, pH 7.4, 21 ºC) showing comparable α-helical content. Secondary structure analysis was conducted with CONTIN/LL [65, 66] at DichroWeb [31 2002, 32, 33] using the reference data set 6 [67], and results can be found in Table 3.

B. Intrinsic tryptophan fluorescence (excitation at 295 nm) indicates that all three isoforms have a similar tertiary structure in phosphate buffer. C. Changes in helical content were followed at 222 nm a Boltzmann sigmoidal equation was fitted to the data, showing the order of stability E2>>E3>=E4 (i).

The fraction unfolded of each protein calculated from the fitted curves was also plotted against temperature (ii). Tm (ºC) for each isoform, the temperature at which there was 50% change in α-
helical content, estimated from the fit denaturation curves can be found in table 4a. Comparable α-helical content at 37°C were also reported in table 4a and fraction unfolded at 37°C is shown in Table 4b. D. Guanidine hydrochloride chemical denaturation of recombinant ApoE isoforms plotted showing the wavelength of maximum fluorescence $\lambda_{\text{max}}$ at an excitation of 292 nm. Experimental data points (closed circles) are shown (i) A three-state unfolding model was fitted to the data; apparent midpoint GuHCl concentrations can be found in Table 4b. Calculated fraction unfolded for each isoform from the fitted three-state unfolding model are shown in (ii). ApoE2 (blue connective line), ApoE3 (green connective line), ApoE4 (red connective line)
Figure 4. Self-assembly of ApoE4 but not ApoE2 or ApoE3

A. Native gel shows the formation of higher oligomeric species for ApoE4 (>1,048 kDa) but not E2 and E3 after incubation at 37°C for 24 h. B. Kinetics of ApoE self-assembly (ApoE2, blue; ApoE3, green; ApoE4, red) was monitored by recording changes in ThT fluorescence intensity at 483 nm for a course of 3 days. A comparison between the three isoforms was established by looking at adjusted ThT fluorescence. Changes in ThT fluorescence varied with each isoform, with a faster increase observed for ApoE4. After 3 days, ApoE3 and ApoE4, but not ApoE2, show changes in ThT fluorescence (top panel). The bottom panel is a close up on the first 24 h of assembly. The length of the lag phase was different for each isoform, with ApoE2>>ApoE3>>ApoE4. Envelopes correspond to
the standard error of the mean (SEM). C. TEM of negatively stained ApoE isoforms after incubation for 24 h (left side) and 3 days (right side) at 37°C show no fibril formation of ApoE2 and ApoE3. ApoE4 self-assembles to form long, curved fibrils. ApoE3 formed round oligomeric species after a 3-day incubation. Scale bars, 500 nm.
Figure 5. Characterisation of ApoE4 fibrils
A. Transmission electron micrographs of negatively stained ApoE4 (25 µM in 20 mM PB, pH 7.4) monitored over 24 h at 37°C. Scale bar, 200 nm. ApoE4 self-assembly was characterised by measuring changes in length (B) and width (C) of the fibrils. Changes in width were non-significant; however, with increasing incubation times, ApoE4 fibrils become significantly longer (average length of 363 nm; One-way ANOVA: ****, p<0.0001, F=180.9). D. CD spectra show retention of α-helical secondary structure after assembly (whole fraction). Fibrils in the pellet were separated from the supernatant to confirm that their secondary structure is not masked by protein in the supernatant. Fibrils in the pellet showed an α-helical conformation. E. X-ray fibre diffraction pattern obtained from partially aligned ApoE4 fibrils after 24 h incubation at 37°C showing positions of diffraction signals on the meridian (vertical) and equatorial (horizontal) axes.
Figure 6. ApoE3 inhibits ApoE4 fibril formation

A. Adjusted ThT fluorescence at 483 nm was monitored over the course of 24 h, 37°C for ApoE3 alone (25 μM), ApoE4 alone (25 μM) and ApoE3 plus ApoE4 (both 12.5 μM for a total ApoE concentration of 25 μM) in PB pH 7.4. B. Transmission electron micrographs showed a heterogeneous population, with the presence of small round and small amorphous species, but no mature fibrils when E4 is incubated with E3 in a 1:1 ratio. Scale bar, 500 nm.
### ApoE2, ApoE3 & ApoE4

<table>
<thead>
<tr>
<th>Parameter</th>
<th>ApoE2</th>
<th>ApoE3</th>
<th>ApoE4</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Gel filtration</strong>*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mw [kDa]</td>
<td>333</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rs [nm]</td>
<td>5.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$f/f_0 = R_s/R_{min}$</td>
<td>1.71</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$D \ [cm^2 \ s^{-1}] \times 10^7$</td>
<td>2.47</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>SEC MALS</strong>*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mw [kDa]</td>
<td>144</td>
<td>139</td>
<td>151</td>
</tr>
<tr>
<td>Rs [nm]</td>
<td>6.55</td>
<td>6.4</td>
<td>6.81</td>
</tr>
<tr>
<td>$f/f_0 = R_s/R_{min}$</td>
<td>1.93</td>
<td>1.88</td>
<td>2.00</td>
</tr>
<tr>
<td>$D \ [cm^2 \ s^{-1}] \times 10^7$</td>
<td>3.78</td>
<td>3.88</td>
<td>3.64</td>
</tr>
</tbody>
</table>

*Values calculated at peak elution volume.

---

**Table 1. Parameters evaluated by gel filtration and SEC MALS**
### Table 2a. Sedimentation analysis of recombinant ApoE isoforms

<table>
<thead>
<tr>
<th>Buffer</th>
<th>Friction ratio†</th>
<th>Sedimentation velocity (S)*</th>
<th>Estimated Molecular Weight*</th>
<th>Peak integration*</th>
<th>Overall RMSD value ±</th>
</tr>
</thead>
<tbody>
<tr>
<td>SEC</td>
<td>ApoE2 1.75</td>
<td>3.11</td>
<td>122 kDa</td>
<td>~90.7%</td>
<td>0.003370</td>
</tr>
<tr>
<td></td>
<td>ApoE3 1.77</td>
<td>3.26</td>
<td>129 kDa</td>
<td>~70.3%</td>
<td>0.003085</td>
</tr>
<tr>
<td></td>
<td>ApoE4 1.8</td>
<td>3.11</td>
<td>127 kDa</td>
<td>~89.5%</td>
<td>0.002772</td>
</tr>
<tr>
<td>PB</td>
<td>ApoE2 1.77</td>
<td>5.07</td>
<td>127 kDa</td>
<td>~84%</td>
<td>0.003275</td>
</tr>
<tr>
<td></td>
<td>ApoE3 1.81</td>
<td>5.10</td>
<td>128 kDa</td>
<td>~88%</td>
<td>0.002749</td>
</tr>
<tr>
<td></td>
<td>ApoE4 1.74</td>
<td>5.23</td>
<td>130 kDa</td>
<td>~86.5%</td>
<td>0.002789</td>
</tr>
</tbody>
</table>

*Corresponding to the main sedimentation peak; † Best fitted friction ratio used to calculate continuous size distribution; ± The root-mean-square deviation (RMSD) evidences the goodness of the fit. Sedimentation analysis of recombinant ApoE isoforms shows that all three ApoE isoforms exist in a tetrameric form in size exclusion buffer (SEC, 20 mM HEPES, 300 mM NaCl, 10 % (v/v) glycerol), pH 8.0 and in 20 mM phosphate buffer (PB), pH 7.4 according to their estimated molecular weight (MW) in solution.
<table>
<thead>
<tr>
<th></th>
<th>ApoE2</th>
<th>ApoE3</th>
<th>ApoE4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ATSAS</td>
<td>ScAtter</td>
<td>ATSAS</td>
</tr>
<tr>
<td><strong>Gunier analysis</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gunier R&lt;sub&gt;G&lt;/sub&gt; (Å)</td>
<td>56.35 ± 0.15</td>
<td>56.34 ± 0.63</td>
<td>56.69 ± 0.14</td>
</tr>
<tr>
<td></td>
<td>0.25 ±</td>
<td>0.28 ±</td>
<td>0.25 ±</td>
</tr>
<tr>
<td>Gunier I(0), arbitrary units</td>
<td>0.00049</td>
<td>0.00044</td>
<td>0.0004</td>
</tr>
<tr>
<td><strong>P(r) analysis (manual)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P(r) R&lt;sub&gt;G&lt;/sub&gt;, (Å)</td>
<td>58.20 ± 0.08</td>
<td>57.56 ± 1.88</td>
<td>58.56 ± 0.088</td>
</tr>
<tr>
<td></td>
<td>0.28 ±</td>
<td>0.27 ± 0.0094</td>
<td>0.25 ± 0.32</td>
</tr>
<tr>
<td>P(r) I(0), arbitrary units</td>
<td>0.00035</td>
<td>0.000946</td>
<td>0.0004</td>
</tr>
<tr>
<td>D&lt;sub&gt;max&lt;/sub&gt;, (Å)</td>
<td>194.5</td>
<td>194.5</td>
<td>195</td>
</tr>
<tr>
<td>q range, (Å&lt;sup&gt;-1&lt;/sup&gt;)</td>
<td>0.0089 - 0.0083</td>
<td>0.0056 - 0.0054</td>
<td>0.0058 - 0.0056</td>
</tr>
<tr>
<td>GNOM total estimate</td>
<td>0.92</td>
<td>---</td>
<td>0.922</td>
</tr>
<tr>
<td>χ² (S&lt;sub&gt;k2&lt;/sub&gt;)</td>
<td>---</td>
<td>1.17 (0.33)</td>
<td>---</td>
</tr>
<tr>
<td><strong>P(r) analysis (Autognom)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P(r) R&lt;sub&gt;G&lt;/sub&gt;, (Å)</td>
<td>58.48 ± 0.12</td>
<td>---</td>
<td>58.69 ± 0.11</td>
</tr>
<tr>
<td></td>
<td>0.28 ±</td>
<td>---</td>
<td>0.25 ±</td>
</tr>
<tr>
<td>P(r) I(0), arbitrary units</td>
<td>0.00046</td>
<td>0.00037</td>
<td>0.0004</td>
</tr>
<tr>
<td>D&lt;sub&gt;max&lt;/sub&gt;, (Å)</td>
<td>204.68</td>
<td>---</td>
<td>203.65</td>
</tr>
<tr>
<td>q range, (Å&lt;sup&gt;-1&lt;/sup&gt;)</td>
<td>0.0054 - 0.0054</td>
<td>0.0054 - 0.0054</td>
<td>0.0056 - 0.0056</td>
</tr>
<tr>
<td>GNOM total estimate</td>
<td>0.896</td>
<td>---</td>
<td>0.899</td>
</tr>
<tr>
<td><strong>Volume of correlation (V&lt;sub&gt;c&lt;/sub&gt;) and</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M&lt;sub&gt;W&lt;/sub&gt; estimate</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Volume of correlation V&lt;sub&gt;c&lt;/sub&gt;, (Å&lt;sup&gt;2&lt;/sup&gt;)</td>
<td>---</td>
<td>1062</td>
<td>---</td>
</tr>
<tr>
<td>M&lt;sub&gt;W&lt;/sub&gt; estimate (kDa)*</td>
<td>---</td>
<td>159 ± 5</td>
<td>---</td>
</tr>
</tbody>
</table>

**Table 2b. Evaluation of the SAXS data by ATSAS and ScAtter.**

Parameters were calculated for comparison in ATSAS 2.8.2 and ScAtter 3.1R; † P(r) based V<sub>c</sub>; * Mass estimates based on P(r) R<sub>G</sub> and V<sub>c</sub> (2).
<table>
<thead>
<tr>
<th></th>
<th>α-Helix</th>
<th>β-Strand</th>
<th>β-Turn</th>
<th>Random Coil</th>
<th>NRMSD</th>
</tr>
</thead>
<tbody>
<tr>
<td>ApoE2</td>
<td>57.3%</td>
<td>3.5%</td>
<td>14.1%</td>
<td>25.3%</td>
<td>0.0178</td>
</tr>
<tr>
<td>ApoE3</td>
<td>59.5%</td>
<td>3.3%</td>
<td>13.0%</td>
<td>24.2%</td>
<td>0.0190</td>
</tr>
<tr>
<td>ApoE4</td>
<td>58.4%</td>
<td>3.4%</td>
<td>13.5%</td>
<td>24.7%</td>
<td>0.0184</td>
</tr>
</tbody>
</table>

Table 3. Secondary structure analysis of recombinant ApoE isoforms from CD spectroscopy

Data

*Calculated by CONTIN programme at DichroWeb (http://dichroweb.cryst.bbk.ac.uk/html/home.shtml)

The normalized root mean square deviation (NRMSD) is an estimate of the goodness-of-fit (NRMSD<0.1 is a necessary but not sufficient condition).

α-helical content was similar between the three ApoE isoforms (within experimental error).
<table>
<thead>
<tr>
<th></th>
<th>Tm (°C)</th>
<th>[Θ]$_{222}$ nm at 37°C (deg.cm$^2$.dmol$^-1$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ApoE2</td>
<td>60.28 ± 1.18</td>
<td>-19181 ± 649</td>
</tr>
<tr>
<td>ApoE3</td>
<td>52.39 ± 1.92</td>
<td>-19809 ± 961</td>
</tr>
<tr>
<td>ApoE4</td>
<td>51.32 ± 4.32</td>
<td>-19092 ± 773</td>
</tr>
</tbody>
</table>

**Table 4a. Thermal denaturation parameters from Boltzmann sigmoidal equation curve**

Tm (°C) represents the temperature at which there was 50% change in α-helical content.

Non-linear best-fit parameters were determined on GraphPad Prism.

<table>
<thead>
<tr>
<th></th>
<th>Fraction unfolded at 37°C</th>
<th>% Fraction unfolded</th>
</tr>
</thead>
<tbody>
<tr>
<td>ApoE2</td>
<td>0.1 ± 0.07</td>
<td>10%</td>
</tr>
<tr>
<td>ApoE3</td>
<td>0.13 ± 0.03</td>
<td>13%</td>
</tr>
<tr>
<td>ApoE4</td>
<td>0.17 ± 0.06</td>
<td>17%</td>
</tr>
</tbody>
</table>

**Table 4b. Fraction unfolded of ApoE at 37°C**

Fraction unfolded of ApoE at 37°C were calculated from the fraction unfolded curves given by equation 7 (Figure 3c). Non-linear best-fit parameters were determined using GraphPad Prism.
<table>
<thead>
<tr>
<th></th>
<th>[GuHCl]$_{50,I}$ (M)</th>
<th>[GuHCl]$_{50,U}$ (M)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ApoE2</td>
<td>0.69 ± 0.03</td>
<td>2.84 ± 0.04</td>
</tr>
<tr>
<td>ApoE3</td>
<td>0.71 ± 0.03</td>
<td>2.66 ± 0.07</td>
</tr>
<tr>
<td>ApoE4</td>
<td>0.58 ± 0.04</td>
<td>2.65 ± 0.08</td>
</tr>
</tbody>
</table>

Table 4c. Chemical denaturation parameters from double sigmoidal curve

[GuHCl]$_{50,I}$ is the concentration needed for 50% of the protein to unfold to an intermediate state. [GuHCl]$_{50,U}$ is concentration needed for 50% of the intermediate state to reach the unfolded state. Non-linear best-fit parameters were determined using GraphPad Prism.
The molecular basis for Apolipoprotein E4 as the major risk factor for late onset Alzheimer’s disease

Ana-Caroline Raulin*, Lucas Kraft*, Youssra K. Al-Hilaly1, Wei-Feng Xue3, John E. McGeehan4, John R. Atack2a, Louise Serpell1x

*authors contributed equally to this manuscript

1. Sussex Neuroscience, School of Life Sciences, University of Sussex, Falmer, Brighton, East Sussex, BN1 6NN, UK
2. Sussex Drug Discovery Centre, School of Life Sciences, University of Sussex, Falmer, Brighton, East Sussex, BN1 6NN, UK
3. School of Biosciences, University of Kent, CT2 7NJ
4. School of Biological Sciences, Institute of Biological and Biomedical Sciences, Faculty of Science, University of Portsmouth, Portsmouth, Hampshire, PO1 2DY, UK

*a Current address: Medicines Discovery Institute, Cardiff University, Cardiff, CF10 3AT, UK

x corresponding author

Running title: Structural basis for increased AD risk in ApoE4 genotype
Abstract

Apolipoprotein E4 (ApoE4) is one of three (E2, E3 and E4) human isoforms of an $\alpha$-helical, 299-amino acid protein. Homozygosity for the $\varepsilon$4 allele is the major risk factor for developing late onset Alzheimer’s disease (AD). ApoE2, ApoE3 and ApoE4 differ at amino acid positions 112 and 158 and these sequence variations may confer conformational differences that underlie their participation in the risk of developing AD. Here, we compared the shape, oligomerisation state, conformation and stability of ApoE isoforms using a range of complementary biophysical methods including small angle X-ray scattering, analytical ultracentrifugation, circular dichroism, X-ray fibre diffraction and transmission electron microscopy. We provide an in-depth and definitive study demonstrating that all three proteins are similar in stability and conformation. However, we show that ApoE4 has a propensity to polymerise to form wavy filaments which do not share the characteristics of cross-$\beta$ amyloid fibrils. Moreover, we provide evidence for the inhibition of ApoE4 fibril formation by ApoE3. This study shows that recombinant ApoE isoforms show no significant differences at the structural or conformational level. However, self-assembly of the ApoE4 isoform may play a role in pathogenesis and these results open opportunities for uncovering new triggers for AD onset.

Key words: Apolipoprotein E; Alzheimer’s disease, small-angle X-ray scattering, analytical ultracentrifugation, alpha-helix
Background

Alzheimer’s disease (AD) is the most prevalent dementia, and the sporadic, late onset form comprises 95% of all AD cases. Diagnosis is based on the observation of specific brain pathology of extracellular amyloid plaques composed of Amyloid-beta (Aβ) peptide and intracellular neurofibrillary tangles formed by tau [1]. The biggest risk factor for developing AD is age, but the ε4 variant of the apolipoprotein E (APOE) gene is the strongest genetic risk factor for the development of late onset AD. The human APOE gene encodes three major protein isoforms, ApoE2, ApoE3 or ApoE4 that differ from one another at amino acid positions 112 and 158. While E2 contains two cysteine residues, E3 contains a cysteine and an arginine, and E4 contains two arginine residues at both sites respectively [2]. ApoE4 presents a risk in a gene dose-dependent manner, with ε3/ε4 heterozygotes having a risk increased by three times and ε4/ε4 homozygotes having up to 15 times more chance of developing AD [3]. ApoE3 is the most common isoform, associated with a neutral risk. ApoE2 is under-represented in the population but is thought to be associated with a lower propensity for AD [4].

ApoE is a predominantly α-helical protein with 299 amino acids that is mainly produced in the liver and by astrocytes in the brain, as well as neurons particularly under stress conditions [5, 6]. ApoE is a component of lipoprotein particles [7] and associates with cholesterol, triglycerides and phospholipids [8]. It is found in the central nervous system participating in high-density lipoproteins (HDL) and in plasma associated with very low-density lipoproteins (VLDL), chylomicron remnants and HDL [9]. It has been suggested that the ApoE isoforms have different lipid binding preferences and that lipid binding may be mediated by a conformational change of the monomer [10]. The ApoE isoform amino acid substitutions are presumed to affect ApoE structure and function and have been suggested to be responsible for ApoE4 homozygotes increased risk for AD [11]. However, the mechanism(s) whereby ApoE4 confers an increased risk of developing AD, whether by a loss-of-function or a gain-of-toxic-function, remain poorly understood. It has been shown that ApoE forms complexes with Aβ through regions within their heparin-binding sites in vitro [12] and was found to be associated with Aβ senile plaques in human brain [13] as well as with neurofibrillary tangles [14]. Histopathological
examination of post-mortem AD brains found a positive correlation between plaque density and
dosage of the ε4 allele [14-16] and positron emission tomography studies matched these findings with
ε4 carriers having higher Aβ deposition compared to non-carriers [17, 18]. ApoE4 has therefore been
suggested to enhance Aβ’s fibrillization and reduce Aβ clearance which is thought to contribute to
AD progression [19]. Especially, aggregated forms of ApoE4 seem to enhance Aβ fibrillization and
antibodies that specifically recognize non-lipidated and aggregated ApoE have been shown to reduce
Aβ deposition in APPPS1-21/APOE4 mice [20]. Recent work has also shown that human ApoE4
transgenic animals have increased deposition of tau compared to ApoE3 human transgenics [21]
indicating a link to accumulation of neurofibrillary tangles. Additionally, human induced pluripotent
stem cell derived neurons expressing ApoE4 were shown to have higher levels of tau phosphorylation
as well as increased degeneration of GABAergic neurons that could be rescued by gene editing and
converting ApoE4 into ApoE3 [22].

The NMR structure of full-length ApoE3 reveals a large, predominantly α-helical globular protein,
with a flexible C-terminal region composed of a short helix (270-277) followed by an unstructured
tail (residues 277-299). The structures for N-terminal domains for ApoE2 and ApoE3 have been
solved by X-ray crystallography [23, 24] and show α-helical bundles which map closely to the N-
terminal domain of E3. It has been proposed that the amino acid substitutions lead to lowered thermal
and chemical stability of ApoE4 and higher propensity to form stable intermediate unfolding states
characteristic of molten globules arising from the formation of a salt-bridge that links the C and N
terminal domain in ApoE4 [25, 26]. A potential consequence of this instability would be an enhanced
proteolysis of ApoE4 in neurons, generating fragments that may interfere with cytoskeletal
components such as tau protein and neurofilaments [27 206, 28]. Decreased stability and molten
globule formation has additionally been suggested to drive ApoE4 to form toxic, fibril-like oligomers
from full length protein, whereas ApoE2 and ApoE3 do not appear to form high molecular weight
species [29].
Here we conducted an in-depth comparison of the structure and stability of the three recombinantly expressed and purified ApoE isoforms using state-of-the-art biophysical techniques and explored ApoE4’s intrinsic property to oligomerise and aggregate. In contrast to previous studies [25, 26], we show that the three recombinant isoforms share very similar quaternary, tertiary and secondary structures and thermal and chemical stability. We reveal that recombinant ApoE4 forms fibrillar structures, however the resulting curvilinear fibrils are fragile and do not share the typical characteristics of cross-β amyloid. Furthermore, to investigate the ApoE4 dose-effect, we explored the effect of co-assembly of ApoE3 and E4. The presence of ApoE3 resulted decreased rate of ApoE4 self-assembly and an inhibition of elongation to long filamentous structures.

Results

**All three ApoE isoforms show similar size and shape**

We have developed a method to produce and maintain full length ApoE proteins recombinantly (see methods and [30]) and it was therefore necessary to first fully characterise these proteins. Initial purification confirmed that ApoE2, E3 and E4 have identical mobilities of 34 kDa by non-reducing SDS PAGE (Figure S1). Gel filtration and SEC coupled multi-angle light scattering (SEC-MALS) showed multimerization of all three proteins consistent with a tetramer in solution with a molecular weight of 139-151 kDa (Figure 1A & B, Table 1). A hydrodynamic radius of 5.8 – 6.5 nm was calculated for each isoform and a frictional ratio f/f₀ above 1.7 suggests elongated shape for the homotetramers (Table 1, Figure S2). Analytical ultracentrifugation was used to investigate the size and shape in solution in further detail. As we extensively dialyzed or buffer exchanged ApoE isoforms into phosphate buffer prior to stability and aggregation studies, we were interested if dialysis affected ApoE oligomerisation in solution. We therefore compared sedimentation in the original size exclusion buffer and in phosphate buffer after dialysis. AUC revealed no differences in sedimentation
velocity between the three isoforms in either buffer (Figure S3, Figure 1 C and D, Table 2a) and the major species was characterised with a molecular mass of 130 kDa (Table 2a).

Small angle X-ray scattering (SAXS) was used to explore whether the tetramers of the ApoE isoforms differ in dimensions and shape. All three proteins exhibited identical scattering profiles indicating that they are very similar in shape (Figure 2A, B, C). A radius of Gyration ($R_g$) of 5.6 nm (56 Å) was calculated using the Guinier approximation and comes close to the hydrodynamic radius determined in the other techniques (Table 2b). The Kratky plot indicates that ApoE consist of several domains that are tethered by linkers with extended conformation. This elongated shape is reproduced in the pair distance distribution function and fits to a maximal dimension of approximately 19.5 nm (195 Å) (Figure 2D).

Together AUC, SEC-MALS and SAXS show that the three ApoE isoforms are similar in size, shape and multimerization, and all form tetrameric species in solution.

**ApoE2, E3 and E4 conformationally similar and show only marginal differences in stability**

Circular dichroism (CD) spectroscopy and tryptophan fluorescence were used to probe potential differences in secondary and tertiary structure of the three isoforms. Negligible spectral differences were observed, suggesting no major conformational differences between the three proteins (Figure 3A and B). Analysis of CD spectra using Dichroweb [31 20 02, 32, 33] showed that the three isoforms all possessed around 58% α-helical content (Table 3), with no significant differences between them (one-way ANOVA: F(2,9)=4.197, p>0.05).

Folding stability of ApoE proteins was examined using thermal and chemical denaturation and monitored using CD and tryptophan fluorescence. Thermal denaturation monitored using CD at 222nm showed a sigmoidal unfolding curve for all three isoforms. A phenomenological Boltzman sigmoidal curve was fitted to the data and revealed that ApoE3 and ApoE4 have similar melting temperature of 52.39°C and 51.32°C respectively, compared to the apparent higher melting
temperature for ApoE2 of 60.28°C (Figure 3C(i)). However, despite these slight differences in the apparent mid-point melting temperature, no differences were observed at physiological temperature of 37°C. One-way ANOVA provided evidence of no significant difference in the mean ellipticity at 222 nm between the three isoforms at 37°C (F(2,9)= 0.9426, p=0.4249; Table 4a). The curves showing fraction unfolded protein for each isoform calculated based on the fitted sigmoidal curves allow direct comparison of the estimated temperature unfolding transitions ((Figure 3C(ii); Figure S4 A), showing similar thermal denaturation profile for all three isoforms. Furthermore, the fraction unfolded at 37°C showed there were no significant differences between the three isoforms at physiological temperature, consistent with the mean ellipticity at 222 nm at the same temperature (Table 4b).

Chemical denaturation in guanidium hydrochloride (GuHCl) was performed in reducing conditions (1 mM DTT) and the wavelength of maximum fluorescence emission intensity $\lambda_{\text{max}}$ was plotted against corresponding concentrations of GuHCl. All three denaturation curves demonstrate multi-phasic behaviour and a two-state unfolding model did not fit the denaturation data. Instead, a three-state unfolding model was fitted to the data (Figure 3D(i)), suggesting the existence of at least three different states of the protein, and the existence of one or more intermediate states. The data is, therefore, consistent with the separate unfolding of two protein domains. From the fitted curves, GuHCl concentrations corresponding to the midpoint of each transition were calculated (Table 4c). Similar to the thermal denaturation study, fraction unfolded curves were calculated for each isoform in order to directly compare the shape of the chemical denaturation profile (Figure 3D(ii)). The transition from a folded state to an intermediate state occurs at a slightly lower GuHCl concentration for ApoE4, suggesting that the first domain to unfold denatures marginally more easily for ApoE4 than ApoE2 and ApoE3 (0.58 M vs ~0.71 M). ApoE3 and ApoE4 show similar unfolding of the second domain, while ApoE2 is more stable (~2.65 M vs 2.84 M). Overall, the observed increased chemical denaturant concentration and temperature needed for ApoE2 to unfold compared with ApoE3 and E4 infers that ApoE2 has a slight increased resistance to denaturation compared to E3 and E4 although the difference in apparent stability is marginal (Table 4c).
ApoE4 assembles to form non-amyloid, native-like fibres

To further investigate the behaviour of the ApoE isoforms, the three proteins were incubated in 20 mM PB at 37°C for 24 h. Native PAGE showed that ApoE2 and E3 ran consistently at a similar mobility at 0 h and after 24 h incubation. However, ApoE4 also formed higher MW species prior to, and following incubation. By 24 h, the majority of the ApoE4 protein formed higher oligomers that did not run through the gel indicating oligomerisation of the ApoE4 protein (Figure 4A). Thioflavin T (ThT) fluorescence assay is frequently used to monitor molecular self-assembly in solution and although it is often utilised to monitor amyloid formation [34], it is not specific for β-sheet structure or for amyloid [35, 36]. ThT was excited at 440 nm and its fluorescence intensity was followed at 483 nm. The change in fluorescence adjusted to baseline displayed a rapid increase in intensity at 483 nm for ApoE4 (Figure 4B(i&ii)), and an extended lag phase for ApoE3 before any changes were detected (Figure B(ii). There was very little change in fluorescence for ApoE2 up to 24 h and only minor increase up to 60 h (Figure 4B). After 3 days incubation at 37°C, the ThT fluorescence increased further for ApoE4, while no further differences in ThT intensity were observed for ApoE2 (Figure S5). ThT fluorescence data indicate fibrillisation of ApoE4, while only minor changes and no changes were observed for ApoE3 and E2 respectively. TEM was used to observe morphological changes with time. TEM of ApoE2 showed some small amorphous and some round species after 24 h and 3 days incubation respectively (Figure 4C). Small round species were observed at 24 h for ApoE3 that increased in size over time and became slightly elongated after 3 days incubation (Figure 4C). In contrast, ApoE4 showed fibrillar structures with a curvy appearance at 24 h (Figure 4C).

To further explore the assembly of ApoE4, TEM was used to monitor the size and morphology of the assemblies at different time points up to 24 h. The electron micrographs show accumulation of small rounded species after only 1 h incubation. The formation of short, curved filaments is detected by 3 h
incubation, which develop into curvy-linear filaments by 6 h and elongate into extended, curvy
filaments by 24 h (Figure 5A). Measurement of the average length of the filaments (Figure 5B)
confirms that the filaments extend in length while the diameter of the structures remains invariant
over 24 h, with an average length of 363 nm and width of 30 nm (Figure 5B and C). Close inspection
of the filamentous structures shows that they have a granular appearance, and do not share the twisted
and smooth appearance generally observed for amyloid fibrils [37-39]. Fibril formation appears to
arise from the end-to-end fusion of the small spherical species, making this unit the apparent smallest
building block of the structure.

CD and X-ray fibre diffraction experiments were conducted to investigate the nature of ApoE4 fibrils
and to investigate whether these assemblies are amyloid-like in structure and undergo the expected
conformational change to β-sheet [29]. CD spectra collected for ApoE4 fibrils incubated for 24 h
showed a spectrum with minima at 208 nm and 222 nm consistent with a predominantly α-helical
secondary structure content (Figure 5D), of comparable shape and intensities to that of non-assembled
ApoE4 (Figure 3A). To ensure that any remaining soluble protein does not dominate the spectra, the
sample was centrifuged at high speed to separate supernatant and pellet fractions. The CD data for the
resolubilised pellet fraction show E4 retains its α-helical structure after incubation at 37°C for 24 h,
and that there is almost no soluble protein left given the very small CD signal in the supernatant
fraction (Figure 5D). The fibre pellet was sonicated to ensure that protein was resuspended
sufficiently and the spectrum continued to demonstrate α-helical content.

To examine the molecular structure of the mature ApoE4 filaments, 100 µM ApoE4 was incubated
for 24 h and aligned to form a partially-aligned fibre bundle. The X-ray fibre diffraction pattern
showed a sharp meridional diffraction signal at 4.5 Å with a weaker diffraction signal at 4.1 Å. On the
equator, a diffuse, strong signal was observed at 10.1 Å and a weaker reflection at approximately 21
Å (Figure 5E). Comparison of the relative intensities shows that the meridional reflections are weaker
than those on the equator which is dissimilar to those usually observed for amyloid fibril patterns
where signals arise from cross-β structure at 4.7 Å and 10 Å on perpendicular axes [40]. The diffraction data obtained supports the CD data described above showing that the ApoE4 retains an α-helical conformation in the fibres which is similar to the structure in the soluble protein.

**ApoE3 and ApoE4 influence the assembly of one another**

ApoE4 alone is able to self-assemble to form non-amyloid-like filaments. However, the majority of individuals with an ε4 allele are heterozygous and have an ε3 allele. To investigate the influence of ApoE3 presence on ApoE4 assembly and vice versa, ApoE3 was incubated with ApoE4 over the course of 24 h to investigate whether they could enhance/seed or inhibit assembly. ApoE3 and ApoE4 were each at 12.5 μM to give a combined ApoE concentration of 25 μM. Assembly over time was compared to assembly of ApoE3 or ApoE4 alone at 12.5 μM (not shown) or 25 μM. Fig 6A shows that ApoE3 reduced the rate of assembly of ApoE4, measured by ThT fluorescence; the ThT kinetics of the mixed sample did not resemble that of E3 alone either. TEM micrographs were produced for the mixed sample and showed the presence of small round and small amorphous species instead of the long mature fibrils observed for ApoE4 alone (Figure 6B).

**Discussion**

The manner in which APOE genotype leads to an increased risk for late onset AD remains unclear. Some studies have proposed that the three ApoE isoforms differ in their structure and consequently function, which may explain ApoE4’s differential effects on AD pathogenesis [3, 41]. Modelling and fluorescence resonance energy transfer studies have been used to provide evidence of a more open conformation for ApoE4 compared to ApoE3 [11, 42, 43] suggesting that the N-terminal region of ApoE4 is less well tethered to the C-terminal domain than for ApoE3. Others have pointed to differences in stability and conformation [29, 44] and have suggested that ApoE4 is less stable and more prone to aggregation to form amyloid-like fibrils [29]. In vivo, ApoE is glycosylated and also
lipidated [45]. While we have explored the structural conformations adopted by recombinant non-modified ApoE, previous studies have shown that the recombinant ApoE produced in \textit{E.coli} adopts a very similar conformation and folding to protein produced in adenovirus [46]. Further studies will be necessary to explore the behaviour of lipidated proteins.

Here we have investigated the size, oligomerisation state and shape of recombinant ApoE isoforms in aqueous solution. Our results are in agreement with previous observations [47] and confirm that all three isoforms form an elongated tetramer in solution. SAXS data provides additional information regarding the low-resolution structure of ApoE isoforms in solution and validates the elongated shape in solution with ApoE qualitatively resembling multiple domains that are tethered by flexible linkers. Identical X-ray scattering profiles between ApoE isoforms suggest no major structural differences in macromolecular architecture.

CD reveals that all three isoforms share a similar α-helical content at physiological pH, when measured at 21°C. This is in contrast to other studies, which have reported differences in α-helical content between the isoforms at 15°C, with ApoE4 being the least α-helical protein [44]. Here, temperature and chemical denaturation studies revealed very small differences in stability towards unfolding for the three isoforms. Melting curves displayed a sigmoidal shape that is characteristic for two-state unfolding, with Tm(E2)≥Tm(E3)≈Tm(E4), which is in agreement with previously reported results in terms of order of stability [25, 44, 48]. However, it is also important to highlight that all three isoforms have the very similar secondary structure content and fraction unfolded at physiological temperature of 37°C.

Chemical denaturation displayed multi-phasic unfolding curves for all three isoforms, consistent with the independent unfolding of the C-terminal and N-terminal domains as described by Morrow and colleagues [25]. While the curves cannot be used to fully deduce $\Delta G^\circ_\text{H}_2\text{O}$ and the slope m-value for each transition since the three-state model employed is empirical in this case, we interpret slight differences in the apparent denaturant concentration corresponding to the midpoint of each transitions.
On the one hand, the transition from the folded to intermediate state was very similar for ApoE2 and ApoE3, however, it occurred at a lower denaturant concentration for ApoE4. On the other hand, transitioning from intermediate to unfolded was most similar between ApoE3 and ApoE4, whereas ApoE2 required a slightly higher denaturant concentration. Others have gone further with chemical denaturation studies by looking at the 10 kDa carboxy- and the 22 kDa amino-terminal domains, in conjunction to the corresponding full-length proteins. Data resulted in the attribution of the first transition to the unfolding of the 10 kDa carboxy fragment and the second to that of the 22 kDa amino-terminal domain, whilst reporting the same order of chemical stability [25, 44, 49].

The loss of stability of ApoE4 as measured by spectroscopic methods, albeit under extreme conditions such as high temperature or under denaturing conditions, may be related to its propensity to polymerise into filamentous structures. We showed here that ApoE4 forms filamentous structures after only 24 h incubation at physiological temperature, pH 7.4, while ApoE2 and ApoE3 remained soluble and globular under the same conditions. TEM reveals that the ApoE4 filaments have a polymeric appearance (beads on a string) and there is a clear hierarchical assembly of small spherical species to small, elongated fibrils and finally fibrils, which show identical diameters. CD shows that following incubation, the ApoE4 fibres retain their α-helical secondary structural content. The CD for the whole fraction is almost identical to ApoE4 prior to incubation. To probe whether the α-helical intensity arises from soluble, unassembled protein, we examined the CD from a sedimented sample. Both pellet and supernatant fractions show a spectrum consistent with high α-helical content. Furthermore, the intensity of the spectrum in the supernatant was very low suggesting that the majority of the ApoE4 protein is found within the fibre containing pellet. This result contrasts with a small shift from α-helix to β-sheet previously described by Hatters et al [29]. Thioflavine T fluorescence assay shows increasing intensity at 483 nm with time showing that the fibril formation course can be followed using this chemical rotamer [36, 50]. Although ThT is well characterised as fluorescent probe for amyloid fibril formation, it is a molecular rotor which is not specifically sensitive to β-sheet structures. Amyloid binding studies have revealed that ThT may associate with
aromatic groups including tyrosine ladders [51, 52]. Fluorescence intensity can increase in the
presence of many polymeric molecules and in viscous solutions [51]. Furthermore, ThT is fluorescent
in the presence of cross-α fibrous molecules revealing that its fluorescence does not rely on β-sheet
structure [53].

X-ray fibre diffraction has been classed as a definitive diagnostic for amyloid fibrils, giving the
classical amyloid cross-β diffraction pattern composed of a 4.76 Å meridional arising from hydrogen
bonding β-strands running perpendicular to the fibre axis and a more diffuse equatorial signal that
arises from β-sheet packing that accommodates side chains [54-56]. Here we revealed that the ApoE4
filaments do not give a cross-β pattern, instead diffraction shows a pattern that can be interpreted as α-
helical rich polymer, consistent with the TEM and CD data. An α-helix has a repeat of 5.4 Å parallel
to the axis of the helix (1.5 Å rise per residue, with 3.6 residues/turn) [57]. However, the structure of
the ApoE molecule is not arranged with all helices arranged in a vertical array. 4.5 Å is a repeating
unit along an extended polypeptide chain and was previously observed for dried peptide crystals [58],
whilst 10.1 Å approximates to the packing of helices in the vertically aligned ApoE structure.

Furthermore, amyloid fibrils are characterised by their high stability and resistance to degradation.
For example, amyloid formed by the Alzheimer’s related peptide is SDS resistant [59]. Here we have
shown that the ApoE4 fibrils are soluble and the ApoE4 species run on a native gel at similar mobility
to ApoE3 and ApoE2. Many examples of native like polymers exist including pathologically related
serpins such as α-1-antitrypsin [60] and of course, functional polymeric proteins such as microtubules
and actin.

The majority of individuals are heterozygous ε3/ε4, and this confers an increased risk for AD
development, albeit lower than for ε4/ε4 individuals [61]. Here we investigated the aggregation
potential for mixed population of ApoE3 and ApoE4 and revealed that E3 inhibits E4 fibril formation
in vitro, leading to a mid-range aggregation kinetic for assembly. Therefore the increased dose of
ApoE4 in ε4/ε4 individuals leads to increased assembly compared to half dose of ApoE4 in ε3/ε4 people.

Conclusions
The detailed studies conducted have revealed that recombinant ApoE2, ApoE3 and ApoE4 resemble one another at a quaternary, tertiary and secondary structural level. All three ApoE isoforms form elongated tetramers in solution and each monomer is rich in α-helical conformation. There is no evidence that the aqueous proteins differ at a structural level. Furthermore, chemical and thermal denaturation studies reveal that the three proteins are similarly stable and follow comparable unfolding mechanisms by melting and by chemical denaturation. However, we have shown that ApoE4 has a higher propensity for polymerisation at physiological temperature and pH, to form elongated fibrous structures which retain the native α-helical conformation. The propensity of ApoE4 to self-assemble may play an important role in the mechanism by which it increases susceptibility to develop AD which impacts on ε4/ε4 individuals more severely than those with ε3/ε4.

Materials and Methods
All materials were purchased from Sigma-Aldrich or Fisher Scientific at the highest purity available.

Protein Production
ApoE2, ApoE3 and ApoE4 were cloned, expressed and purified using methods detailed in the supplementary information. Briefly, a codon optimized ApoE4 gene was cloned into a pET17b vector with a six-histidine tag, thioredoxin (TRX) and HRV 3C protease cleavage site upstream. ApoE2 and ApoE3 genes were generated by site directed mutagenesis using QuikChange. Recombinant ApoE proteins were expressed in E. coli Rosetta2(DE3) cells and affinity purified using Talon ® beads followed by HiTrap heparin affinity column and cleaved on column with PreScission protease. The ApoE proteins were eluted using a salt gradient and further purified using a HiLoad Superdex 26/600 200 pg size exclusion column. ApoE containing fractions were concentrated and stored at -80°C.
Gel filtration studies

The high molecular weight protein standard (GE Healthcare, # 28-4038-42) was used to calibrate the HiLoad Superdex 26/600 200 pg column (GE Healthcare). All protein standards and dextran blue were dissolved in size exclusion chromatography (SEC) buffer (20 mM HEPES, 300 mM NaCl, 10% (v/v) glycerol, pH 8.0) to the manufacturers recommended concentrations (0.4 mg/mL Ferritin, 4 mg/mL Aldolase, 3 mg/mL Conalbumin, 3 mg/mL Ovalbumin) and 1.6 mL of the protein calibration mixture applied onto the column at a flow rate of 1.5 mL/min. Stokes radius (size) values were calculated on the basis of ApoE elution volume (equation (1))

\[ K_{av} = \frac{V_e - V_0}{V_t - V_0} \quad (1) \]

where \( V_e \) = elution volume for the protein, \( V_0 \) = Superdex 200 void volume (112.26 mL), and \( V_t \) = total Superdex 200 bed volume (320 mL). \( K_{av} \) values were calculated for each protein standard and a standard curve generated by plotting the \( \sqrt{(-\log(K_{av}))} \) as y values versus the Stoke radius of each standard as x values (Ferritin 6.1 nm, Aldolase 4.8 nm, Conalbumin 4.04 nm, Ovalbumin 2.75 nm). ApoE Stoke radius \( R_S \) was estimated with the following best fit linear equation (2):

\[ y = 0.1058 x + 0.2397 \quad (R^2 = 0.9552) \quad (2) \]

ApoE diffusion coefficient \( D \) was estimated from its Stoke radius \( R_S \) by following equation (3):

\[ D = \frac{kT}{6\pi\eta R_S} \quad (3) \]

where \( k \) is the Boltzmann constant (1.38 x 10^{-23} \text{ m}^2 \text{ kg s}^{-2} \text{ K}^{-1}), \( T \) is the experimental temperature (277.15 K) and \( \eta \) is the solvent viscosity (1.4181 x 10^{-3} \text{ Pa s}) at temperature \( T \).

The frictional ratio \( f/f_0 = R_S/R_{\text{min}} \) was calculated by assuming the minimal radius (\( R_{\text{min}} \)) of a sphere that could contain the given mass of tetrameric ApoE (136800 Da) where \( R_{\text{min}} \) [62] is defined as

\[ R_{\text{min}} = 0.066 M^{\frac{1}{3}} \quad (\text{for } M \text{ in Dalton, } R_{\text{min}} \text{ in nm}) \quad (4) \]

Dialysis and buffer exchange
ApoE was either extensively dialysed overnight in 20 mM phosphate buffer (PB) (16 mM Na₂HPO₄, 4 mM NaH₂PO₄), pH 7.4 using Slide-A-Lyzer™ Dialysis Cassettes with a molecular weight cut-off (MWCO) of at 3.5 kDa (Thermo Fisher) or buffer exchanged using disposable Vivaspin® 500 centrifugal concentrators with a MWCO of 3 kDa or Vivaspin® 20 centrifugal concentrators with a MWCO of 3 kDa or 10 kDa (Sartorius).

**Gel Electrophoresis**

Protein samples for denaturing one dimensional sodium dodecylsulphate (1D-SDS) – gel electrophoresis were prepared in Laemmli sample buffer (Bio-Rad) to a final concentration of 3 µM, omitting the addition of any reducing agent. Samples were loaded into a 4-20% TGX gel (Bio-Rad), and the gel was run at 120 V (constant voltage) for 80 min. Resolved proteins were stained using a Coomassie dye R-250 containing solution (Imperial protein stain, #24615, Thermo Scientific).

Protein samples for native PAGE were diluted to 3 µM in native sample buffer (Bio-Rad) and were directly loaded into a 4-20% TGX gel (Bio-Rad). Gels were run at 120V for 120 min. Resolved proteins were stained using a Coomassie dye R-250 containing solution.

**Small angle X-ray scattering and size exclusion chromatography multi angle light scattering (SEC MALS)**

Small angle X-ray scattering (SAXS) and multi angle light scattering (MALS) experiments were performed at the B21 beamline (Diamond Light Source, UK).

X-ray scattering was acquired with an X-ray wavelength of 1 Å on a Pilatus 2M detector at a distance of 3.9 m and a camera length of 4.036 m. ApoE isoforms at 10 mg/mL were delivered at 20°C and a flow rate of 0.16 mL/min via an in-line Agilent high performance liquid chromatography (HPLC) with a Shodex Kw-403 column and 20 mM HEPES, 300 mM NaCl, pH 8.0 as running buffer. In total 620 frames were recorded and each frame exposed for 3s. Buffer subtraction and averaging was
performed in ScAtter version 3. Data was analysed in ScAtter version 3 and ATSAS 2.8.3 for comparison.

ApoE isoforms were diluted to 5 mg/mL in 20 mM HEPES, 300 mM NaCl, 1 mM TCEP, pH 8.0 for MALS experiments and delivered at RT via an in-line Agilent HPLC with a Shodex Kw-403 column. Refractive increments (dn/dc) were determined by a Wyatt optilab T-rEX and scattering measured by a Wyatt Dawn Heleo with QELS. Data was analysed using ASTRA version 6.1.7 (Wyatt).

Analytical ultracentrifugation

Sedimentation velocity experiments were carried out in a Beckman model XL-A analytical ultracentrifuge (Beckman Coulter, Fullerton, CA) at 20°C using an AnTi60 rotor. ApoE isoforms (400 µL) were used at a concentration of 8 µM in 20 mM HEPES, 300 mM NaCl, 10% (v/v) glycerol, pH 8.0 or at a concentration of 8 µM in 20 mM PB, pH 7.4. Experiments were performed at a rotor speed of 40,000 rpm (128,794 x g) and absorbance at 280 nm was recorded every 25 min, at radial intervals of 0.003 cm. A continuous size distribution c(s) from the Lamm equation model in the range of 0.1S-15S was fit to the data using the programme SEDFIT version 15.01b [63]. The confidence interval (F-ratio) was set to 0.95. Buffer viscosity and density was calculated in Sednterp version 20120828 BETA [64 & Pelletier, 1992].

Circular Dichroism (CD) Spectroscopy

ApoE isoforms were diluted to a final concentration of ~25 µM in 20 mM PB, pH 7.4. The exact concentrations were confirmed in triplicate using a NanoDrop ND2000c spectrophotometer (Thermo Fisher) by measuring the absorbance at 280 nm, and applying Beer-Lambert’s law using an extinction coefficient of 44,460 L.mol⁻¹.cm⁻¹ The averaged concentrations were recorded and used in calculations when needed.

FarUV CD data were collected using a Jasco-715 CD-spectrometer (Jasco, Goh-Umstadt, Germany) at different time points. Temperature was maintained at 21°C using a Peltier controlled cell holder.
All spectra were collected in 0.01 cm transparent quartz cuvettes (Starna Scientific, Essex, UK) in the range of 280-180 nm with a resolution of 0.1 nm and a band-width of 1 nm. Scanning-speed was set to 50 nm/min, response time to 4 s and sensitivity to standard. Three spectra per sample were acquired and averaged to give a spectrum after subtracting blank spectra of the buffer. For each isoform, spectra were collected on protein from at least three different production batches and an average trace of [Θ] against wavelength was obtained using GraphPad Prism. The mean residue ellipticity ([Θ] in deg.cm$^2$.dmol$^{-1}$) was calculated from the measured ellipticity Θ in mdeg at wavelength λ using equation (5):

$$\Theta = \frac{10 \cdot n \cdot C \cdot l}{100 \cdot C \cdot I}$$

with n the number of amino-acid bonds in the protein, C the concentration of the sample in mol.L$^{-1}$ and a path length of 1 cm.

Secondary structure of ApoE isoforms was calculated by deconvolution of circular dichroism spectra performed with CONTIN/LL [65, 66] at DichroWeb [32] using the reference spectra set 6 [67]. Results represent a mean of values from a minimum of three spectra per isoform. One-way ANOVA was performed to compare the α-helical content between the isoforms, using GraphPad Prism.

**Thermal denaturation using CD**

ApoE isoforms were diluted to a final concentration of ~25 μM in 20 mM PB, pH 7.4. CD data were collected using a Jasco-715 CD-spectrometer (Jasco, Goh-Umstadt, Germany). Temperature was increased at a rate of 1°C/min from 20°C to 80°C using a Peltier controlled cell holder. Thermal scans were acquired at a wavelength λ of 222 nm every 1.0°C with a sensitivity of 100 mdeg, a response time of 1s and a bandwidth of 1 nm. Thermal scans were acquired for proteins purified from at least three different production batches and an average trace of [Θ] against temperature was obtained using GraphPad Prism. The unfolding was irreversible. A phenomenological Boltzmann sigmoidal curve with the following equation (6) was fitted to the averaged thermal denaturation curves:
\[ Y = \frac{\text{Top} - \text{Bottom}}{\text{Tm} - X} \]  
(6)

with \( \text{Top} \) = maximum \( \Theta \) value, \( \text{Bottom} \) = minimum \( \Theta \) value, \( \text{Tm} \) the temperature corresponding to 50% change in \( \alpha \)-helical content and \( \text{Slope} \) = the steepness of the curve. The calculated thermal denaturation curve based on the fitted curves was also transformed into the fraction of the protein unfolded for comparison of protein transformation with the following equation (7):

\[
\text{Fraction unfolded} = \frac{[\Theta]_N - [\Theta]}{[\Theta]_N - [\Theta]_D} \quad (7)
\]

with \([\Theta]_N\) the ellipticity of the protein at 20°C corresponding to the folded state and \([\Theta]_D\) the ellipticity of the denatured protein at 80°C.

**Intrinsic fluorescence measurements**

Recombinant ApoE isoforms were diluted to 10 µM in 20 mM PB and incubated overnight at 4°C. Fluorescence scans were acquired using SpectraMax i3 reader (Molecular dimensions). Excitation was set to 295 nm and scans were collected with 1 nm increments between 320 nm and 465 nm. Excitation and emission bandwidths were set to 9 nm and 15 nm respectively. The number of reading per well was set to 6. The photomultiplier tube voltage was set to high. Blank spectra of the buffer were subtracted to protein fluorescence scans. A minimum of 3 different production batches per isoform was used and an average trace of fluorescence intensity against wavelength was obtained using GraphPad Prism.

**Chemical denaturation and data analysis**

Recombinant ApoE isoforms at ~0.05 mg/mL in 20 mM PB, 1 mM dithiothreitol (DTT) were incubated overnight at 4°C with increasing concentrations of guanidine-hydrochloride (GuHCl) (0 to 6 M) [25]. Measurements were obtained with a Varian Cary Eclipse spectrophotometer (Varian Ltd., Oxford, UK), quartz cuvette (1 cm path length; Starna, Essex, UK). Temperature was maintained at 20°C using a Varian Cary temperature controller.
Tryptophan residues were selectively excited at 292 nm and emission was monitored between 310 nm and 400 nm; excitation and emission slits were set to 10 nm. Five acquisitions per concentration point were acquired at a scan speed of 90 nm/min using response time of 0.05 s. A minimum of 3 different production batches per isoform were used [68].

Wavelength of maximum emission ($\lambda_{\text{max}}$) per GuHCl point was determined by peak fitting of the averaged emission spectra [44]. Denaturation curves were obtained by plotting the wavelength corresponding to the maximum fluorescence intensity against GuHCl concentration ([GuHCl]) on GraphPad Prism. The curves show two major transitions and can be described by a three-state unfolding model [69], as described in equation (8) and (9)

$$\text{Folded} \rightleftharpoons \text{Intermediate} \rightleftharpoons \text{Unfolded} \quad (8)$$

$$Y = \frac{Y_F ([\text{GuHCl}]) + Y_U ([\text{GuHCl}]) \cdot K_{\text{app}}}{1 + K_{\text{app}}} \quad (9)$$

Equation (9) [69] was fitted to the data with $Y$ corresponding to $\lambda_{\text{max}}$, and $Y_F$ and $Y_U$ the signal corresponding to the folded (F) and unfolded (U) protein respectively. $Y_F$ and $Y_U$ are assumed to be linearly dependent on the concentration of denaturant, as displayed in equation (10):

$$Y_X = Y_X^0 + m_X [\text{GuHCl}] \quad (10)$$

with $X$ corresponding to either the F or U state and $Y_X^0$ the $\lambda_{\text{max}}$ at either 0 M GuHCl or 6 M. In equation (9) $K_{\text{app}}$ corresponds to the apparent guanidine-dependent equilibrium constant of the denaturation process, defined by equation (11):

$$K_{\text{app}} = \frac{K_1 K_2 + AK_1}{1 + (1 - A) K_1} \quad (11)$$

$K_1$, $K_2$ and $A$ are defined by equation (12-14)

$$K_1 = e^{-\frac{\Delta G_{1H2O}^\circ - m_1 [\text{GuHCl}]}{RT}} \quad (12)$$

$$K_2 = e^{-\frac{\Delta G_{2H2O}^\circ - m_2 [\text{GuHCl}]}{RT}} \quad (13)$$

$$A = \frac{(Y_I - Y_U)}{(Y_P - Y_U)} \quad (14)$$

$\Delta G_{1H2O}^\circ$, $\Delta G_{2H2O}^\circ$, $m_1$ and $m_2$ were the fitted parameters, while $\Delta G_{1H2O}^\circ$ and $\Delta G_{2H2O}^\circ$ are the linearly extrapolated free energy differences at 0 M GuHCl [69] F to I, and I to U, respectively.
Subsequently, the concentrations corresponding to 50% unfolding of the protein to an intermediate state ([GuHCl]_{50,I}) and to an unfolded state ([GuHCl]_{50,U}) were calculated by equations (15) and (16) and compared:

\[
[\text{GuHCl}]_{50,I} = \frac{\Delta G^o_{1\text{H}_2\text{O}}}{m_1} \quad (15)
\]

\[
[\text{GuHCl}]_{50,U} = \frac{\Delta G^o_{2\text{H}_2\text{O}}}{m_2} \quad (16)
\]

The fraction unfolded, \( A \), was also calculated and plotted as a function of [GuHCl] to facilitate comparison of the shape of the unfolding curves between the three isoforms.

**Thioflavin T (ThT) fluorescence assay of ApoE self-assembly**

ApoE isoforms at \(~25 \, \mu\text{M}\) in 20 mM PB, pH 7.4 were incubated in the presence of Thioflavin-T (ThT) at 37°C in the SpectraMax i3 plate reader. Aqueous ThT stock solution was prepared at a concentration of 3.14 mM, filtered through a 0.2 \, \mu\text{m} pore size, and used in a 1:2.2 ratio. 96-well plates were sealed with an optically clear polyolefin film to avoid evaporation (StarSeal Advanced Polyolefin Film, Starlab). The number of reading per well was set to 6, PMT voltage was set to high and blank spectra of the buffer were subtracted to protein fluorescence scans. Excitation wavelength was 440 nm and emission at 483 nm was monitored every 15 min, with 3 s low orbital shakes before readings. Fluorescence intensity at a given time point (F) for each isoform was adjusted by subtracting the minimum fluorescence intensity value (\( F_{\text{min}} \)), and not to the initial fluorescence, to account for the increasing temperature effect on ThT (from room temperature to 37°C).

Adjusted fluorescence was plotted against time, and averaged traces for each isoform was obtained on GraphPad Prism. A minimum of 3 different production batches per isoform was used.

**Negative Stain Transmission Electron Microscopy (TEM)**

Morphology of ApoE after 24 h incubation at 37°C was assessed by negative stain TEM. A droplet of
sample (4 μL) was placed on 400-mesh carbon-coated grids (Agar Scientific, Essex, UK) and incubated for 1 min. After blotting the excess solution, the grid was washed with 4 μL filtered Milli-Q water and blotted. It was then negatively stained with 4 μL filtered 0.5% uranyl acetate for 40 s and blotted with filter paper. Grids were left to air-dry for at least 5 min before storage. Grids were examined on a Jeol Jem1400-plus transmission electron microscope (Jeol, U.S.A) operated at 80 kV fitted with a Gatan Orius SC100 camera (UK).

Separation of fibril pellet from supernatant

To probe the secondary structure of ApoE4 fibrils, pellets were isolated from supernatant by ultracentrifugation of assembled fibrils at 60,000 rpm in an Optima™ MAX Ultracentrifuge (Beckman Coulter) using a Beckman TLA120.2 fixed-angle rotor for 45 min at 4°C. Supernatant was removed and replace by fresh 20 mM PB buffer. CD spectra of the whole fraction, supernatant, pellet and sonicated pellet (5 s sonication) were acquired as described for CD. The data was plotted graphically using GraphPad Prism but without conversion to molar ellipticity. Data for whole fraction, pellet and supernatant for the same sample were compared directly to compare secondary structure and protein content for each fraction.

X-Ray fibre diffraction

ApoE4 (100 μM in PB pH 7.4) was incubated for 24 h at 37°C to form fibrils and then the sample was centrifuged for 40 min at 21,100 rpm at 4°C using a benchtop centrifuge (Mikro 22R, Hettich). The resulting pellet was resuspended in 400 μL filtered milliQ water and this was repeated to remove salts that could interfere with the diffraction pattern. Finally, the pellet was resuspended in 30 μL filtered milliQ water. 10 μL of concentrated fibre sample was suspended between two wax-filled capillaries and incubated in a sealed Petri dish at 4°C to form a partially aligned fibre sample [70].

X-ray diffraction images were collected using a Rigaku rotating anode source (CuKα) and a Saturn CCD+ detector. Partially aligned fibres were placed in the X-ray beam and exposed for 30 s or 60 s at specimen to detector distances of 50 mm and 100 mm. Diffraction patterns were converted to TIFF format using imosflm [71] and analysed using CLEARER [72].
Abbreviations

Alzheimer’s disease  AD
ApolipoproteinE  ApoE
Transmission electron microscopy  TEM
Small angle X-ray scattering  SAXS
Analytical ultracentrifugation  AUC
Low density lipoproteins  LDL
Very low-density lipoproteins  VLDL
Circular dichroism  CD
thioredoxin  TRX
size exclusion chromatography  SEC
multi angle light scattering  MALS
high performance liquid chromatography  HPLC
phosphate buffer  PB
one dimensional sodium dodecylsulphate 1D-SDS)
dithiothreitol  DTT)
guanidine-hydrochloride  GuHCl)
Thioflavine T  ThT

Declarations

Ethics approval and consent to participate
There are no ethical considerations related to this work

Consent for publication
Not applicable

Availability of data and materials
All data generated or analysed during this study are included in this published article and its supplementary information files.

Competing Interest
The authors declare they have no conflicts of interest

Funding
LK is funded by the Alzheimer’s society Doctoral training centre grant awarded to LCS at University of Sussex. AR is funded by Sussex Neuroscience doctoral training centre supported by the University of Sussex. The LCS is supported by Alzheimer’s society and Alzheimer’s research UK.

**Author contributions**

LK and AR conducted the work and wrote the paper. LK developed the protein purification and collected and analysed SAXS data. AR collected and analysed folding and assembly data and analysed the folding results with WFX. Both AR and LK collected the AUC data. JM supervised the AUC data collection and managed the analysis. YA collected the diffraction data. LCS and JA managed the research. LCS wrote the paper. JA edited the paper.

**Acknowledgements**

The authors gratefully acknowledge Dr Antony Oliver (GDSC) for help with SAXS analysis, Dr Pascale Schellenberger for transmission electron microscopy. The authors thank Dr Laura Itzaki for help with data analysis. We thank Diamond Light Source for access to beamline B21 that contributed to the results presented here. We thank Robert Rambo, Nathan Cowieson and Nikul Khunti for support during beamtime. The authors thank Caroline Lelliot for her contribution to ThT data collection optimisation.

**References**


[38] Serpell L. Alzheimer's amyloid fibrils: structure and assembly. Biochim Biophys Acta.


Figure legends

Figure 1. Characterisation of recombinant ApoE2, E3 and E4.

A. The SEC elution pattern of ApoE isoforms on the Superdex 200 column as a function of absorbance at 280 nm versus the elution volume. All ApoE isoforms have a minor elution peak at 131 mL and a major at 151 mL suggesting the presence of different oligomeric species. A hydrodynamic radius of 5.8 nm, as well as a diffusion coefficient of $2.47 \times 10^{-7}$ cm$^2$s$^{-1}$ is calculated by calibration of the Superdex 200 column (inlet) using commercially available protein standards (1 – ferritin, 2 – aldolase, 3 – conalbumin, 4 – ovalbumin). A frictional ratio $f/f_0 = R_s/R_{\text{min}}$ of 1.71 is calculated which suggest moderate elongation of ApoE [62]. B. SEC MALS plotted with the differential refractive index (RIU, line) are shown as a function of elution time for ApoE2, ApoE3 and ApoE4. Insert shows the calculated molecular weight (Mw) across the peak. All ApoE isoforms have identical elution volumes and a molecular weight of 145 kDa, 139 kDa and 151 kDa is calculated at elution peak for ApoE2, ApoE3 and ApoE4 respectively. C. and D. AUC showing continuous c(S) size distributions in size exclusion buffer (20 mM HEPES, 300 mM NaCl, 10% (v/v) glycerol, pH 8.0) (C) and in 20 mM PB pH 7.4 (D). No difference between ApoE isoforms is observed in either buffer condition. A major species with a sedimentation coefficient $S$ at 3 and 5 is observed in size exclusion and PB buffer respectively. The difference in sedimentation coefficient between buffers is due to the presence of glycerol in the storage buffer.

Figure 2. Small angle X-ray scattering. X-ray scattering curves and the dimensionless Kratky plot (ScÅtter) for A. ApoE2, B. ApoE3 and C. ApoE4 are shown, as well as D. their corresponding pair distance distribution function $P(r)$. A-C. All ApoE isoforms have identical scattering profiles and adopt an extended conformation in solution with some intrinsic level of flexibility as assessed by the dimensionless Kratky plot. D. This extended conformation is seen in the $P(r)$ distribution respectively and a maximal dimension $D_{\text{max}}$ of approximately 19.5 nm (195 Å) is determined for each isoform.
Figure 3. Conformation and stability of ApoE isoforms.

A. Far UV CD spectra of ApoE isoforms (25 µM in 20 mM PB, pH 7.4, 21 °C) showing comparable α-helical content. Secondary structure analysis was conducted with CONTIN/LL [65, 66] at DichroWeb [31 2002, 32, 33] using the reference data set 6 [67], and results can be found in Table 3.

B. Intrinsic tryptophan fluorescence (excitation at 295 nm) indicates that all three isoforms have a similar tertiary structure in phosphate buffer. C. Changes in helical content were followed at 222 nm a Boltzmann sigmoidal equation was fitted to the data, showing the order of stability E2>>E3>=E4 (i).

The fraction unfolded of each protein calculated from the fitted curves was also plotted against temperature (ii). Tm (°C) for each isoform, the temperature at which there was 50% change in α-helical content, estimated from the fit denaturation curves can be found in table 4a. Comparable α-helical content at 37°C were also reported in table 4a and fraction unfolded at 37° C is shown in Table 4b. D. Guanidine hydrochloride chemical denaturation of recombinant ApoE isoforms plotted showing the wavelength of maximum fluorescence λ_{max} at an excitation of 292 nm. Experimental data points (closed circles) are shown (i) A three-state unfolding model was fitted to the data; apparent midpoint GuHCl concentrations can be found in Table 4b. Calculated fraction unfolded for each isoform from the fitted three-state unfolding model are shown in (ii). ApoE2 (blue connective line), ApoE3 (green connective line), ApoE4 (red connective line).

Figure 4. Self-assembly of ApoE4 but not ApoE2 or ApoE3

A. Native gel shows the formation of higher oligomeric species for ApoE4 (>1,048 kDa) but not E2 and E3 after incubation at 37°C for 24 h. B. Kinetics of ApoE self-assembly (ApoE2, blue; ApoE3, green; ApoE4, red) was monitored by recording changes in ThT fluorescence intensity at 483 nm for a course of 3 days. A comparison between the three isoforms was established by looking at adjusted ThT fluorescence. Changes in ThT fluorescence varied with each isoform, with a faster increase observed for ApoE4. After 3 days, ApoE3 and ApoE4, but not ApoE2, show changes in ThT fluorescence (top panel). The bottom panel is a close up on the first 24 h of assembly. The length of the lag phase was different for each isoform, with ApoE2>ApoE3>>ApoE4. Envelopes correspond to the standard error of the mean (SEM). C. TEM of negatively stained ApoE isoforms after incubation
for 24 h (left side) and 3 days (right side) at 37°C show no fibril formation of ApoE2 and ApoE3. ApoE4 self-assembles to form long, curved fibrils. ApoE3 formed round oligomeric species after a 3-day incubation. Scale bars, 500 nm.

Figure 5. Characterisation of ApoE4 fibrils

A. Transmission electron micrographs of negatively stained ApoE4 (25 µM in 20 mM PB, pH 7.4) monitored over 24 h at 37°C. Scale bar, 200 nm. ApoE4 self-assembly was characterised by measuring changes in length (B) and width (C) of the fibrils. Changes in width were non-significant; however, with increasing incubation times, ApoE4 fibrils become significantly longer (average length of 363 nm; One-way ANOVA. ****, p<0.0001, F=180.9). D. CD spectra show retention of α-helical secondary structure after assembly (whole fraction). Fibrils in the pellet were separated from the supernatant to confirm that their secondary structure is not masked by protein in the supernatant. Fibrils in the pellet showed an α-helical conformation. E. X-ray fibre diffraction pattern obtained from partially aligned ApoE4 fibrils after 24 h incubation at 37°C showing positions of diffraction signals on the meridian (vertical) and equatorial (horizontal) axes.

Figure 6. ApoE3 inhibits ApoE4 fibril formation

A. Adjusted ThT fluorescence at 483 nm was monitored over the course of 24 h, 37°C for ApoE3 alone (25 µM), ApoE4 alone (25 µM) and ApoE3 plus ApoE4 (both 12.5 µM for a total ApoE concentration of 25 µM) in PB pH 7.4. B. Transmission electron micrographs showed a heterogeneous population, with the presence of small round and small amorphous species, but no mature fibrils when E4 is incubated with E3 in a 1:1 ratio. Scale bar, 500 nm.


<table>
<thead>
<tr>
<th></th>
<th>ApoE2, ApoE3 &amp; ApoE4</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Gel filtration</strong></td>
<td></td>
</tr>
<tr>
<td>Mw [kDa]</td>
<td>333</td>
</tr>
<tr>
<td>Rs [nm]</td>
<td>5.8</td>
</tr>
<tr>
<td>$f/f_0 = R_s/R_{\text{min}}$</td>
<td>1.71</td>
</tr>
<tr>
<td>$D \ [\text{cm}^2 \text{s}^{-1}] \times 10^{-7}$</td>
<td>2.47</td>
</tr>
</tbody>
</table>

| **SEC MALS**           |                       |
| Mw [kDa]               | ApoE2 | ApoE3 | ApoE4 |
| Rs [nm]                | 6.55   | 6.4   | 6.81  |
| $f/f_0 = R_s/R_{\text{min}}$ | 1.93   | 1.88  | 2.00  |
| $D \ [\text{cm}^2 \text{s}^{-1}] \times 10^{-7}$ | 3.78   | 3.88  | 3.64  |

Table 1. Parameters evaluated by gel filtration and SEC MALS

*Values calculated at peak elution volume.
<table>
<thead>
<tr>
<th>Buffer</th>
<th>Friction ratio †</th>
<th>Sedimentation velocity (S)*</th>
<th>Estimated Molecular Weight*</th>
<th>Peak integration*</th>
<th>Overall RMSD value ±</th>
</tr>
</thead>
<tbody>
<tr>
<td>SEC</td>
<td>ApoE2 1.75</td>
<td>3.11</td>
<td>122 kDa</td>
<td>~90.7%</td>
<td>0.003370</td>
</tr>
<tr>
<td></td>
<td>ApoE3 1.77</td>
<td>3.26</td>
<td>129 kDa</td>
<td>~70.3%</td>
<td>0.003085</td>
</tr>
<tr>
<td></td>
<td>ApoE4 1.8</td>
<td>3.11</td>
<td>127 kDa</td>
<td>~89.5%</td>
<td>0.002772</td>
</tr>
<tr>
<td>PB</td>
<td>ApoE2 1.77</td>
<td>5.07</td>
<td>127 kDa</td>
<td>~84%</td>
<td>0.003275</td>
</tr>
<tr>
<td></td>
<td>ApoE3 1.81</td>
<td>5.10</td>
<td>128 kDa</td>
<td>~88%</td>
<td>0.002749</td>
</tr>
<tr>
<td></td>
<td>ApoE4 1.74</td>
<td>5.23</td>
<td>130 kDa</td>
<td>~86.5%</td>
<td>0.002789</td>
</tr>
</tbody>
</table>

Table 2a. Sedimentation analysis of recombinant ApoE isoforms

*Corresponding to the main sedimentation peak; †Best fitted friction ratio used to calculate continuous size distribution; ±The root-mean-square deviation (RMSD) evidences the goodness of the fit. Sedimentation analysis of recombinant ApoE isoforms shows that all three ApoE isoforms exist in a tetrameric form in size exclusion buffer (SEC, 20 mM HEPES, 300 mM NaCl, 10 % (v/v) glycerol), pH 8.0 and in 20 mM phosphate buffer (PB), pH 7.4 according to their estimated molecular weight (MW) in solution.
### Table 2b. Evaluation of the SAXS data by ATSAS and ScAtter.

Parameters were calculated for comparison in ATSAS 2.8.2 and ScAtter 3.1R; \(^*\) P(r) based \(V_c\); \(^*\) Mass estimates based on P(r) \(R_g\) and \(V_c\) (2).

<table>
<thead>
<tr>
<th></th>
<th>ApoE2</th>
<th>ApoE3</th>
<th>ApoE4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ATSAS</td>
<td>ScAtter</td>
<td>ATSAS</td>
</tr>
<tr>
<td>Gunier analysis</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Guinier (R_g) (Å)</td>
<td>56.35 ± 0.15</td>
<td>56.34 ± 0.63</td>
<td>56.69 ± 0.14</td>
</tr>
<tr>
<td></td>
<td>0.25 ±</td>
<td>0.28 ±</td>
<td>0.25 ±</td>
</tr>
<tr>
<td>Guinier I(0), arbitrary units</td>
<td>0.00049</td>
<td>0.00044</td>
<td>0.0004</td>
</tr>
<tr>
<td>P(r) analysis (manual)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P(r) (R_g) (Å)</td>
<td>58.20 ± 0.08</td>
<td>57.56 ± 1.88</td>
<td>58.56 ±</td>
</tr>
<tr>
<td></td>
<td>0.28 ±</td>
<td>0.27 ± 0.0094</td>
<td>0.25 ± 0.32</td>
</tr>
<tr>
<td>P(r) I(0), arbitrary units</td>
<td>0.00035</td>
<td>0.00037</td>
<td>0.00034</td>
</tr>
<tr>
<td>(D_{max}) (Å)</td>
<td>194.5</td>
<td>194.5</td>
<td>195</td>
</tr>
<tr>
<td>q range, (Å(^{-1}))</td>
<td>0.0089</td>
<td>0.0083</td>
<td>0.0056</td>
</tr>
<tr>
<td></td>
<td>0.2513</td>
<td>0.2461</td>
<td>0.2414</td>
</tr>
<tr>
<td>GNOM total estimate</td>
<td>0.92</td>
<td>⎯</td>
<td>0.922</td>
</tr>
<tr>
<td>(\chi^2 (S_k^2))</td>
<td>⎯</td>
<td>1.17 (0.33)</td>
<td>⎯</td>
</tr>
<tr>
<td>P(r) analysis (Autognom)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P(r) (R_g) (Å)</td>
<td>58.48 ± 0.12</td>
<td>⎯</td>
<td>58.69 ± 0.11</td>
</tr>
<tr>
<td></td>
<td>0.28 ±</td>
<td>⎯</td>
<td>0.25 ±</td>
</tr>
<tr>
<td>P(r) I(0), arbitrary units</td>
<td>0.00046</td>
<td>⎯</td>
<td>0.00037</td>
</tr>
<tr>
<td>(D_{max}) (Å)</td>
<td>204.68</td>
<td>⎯</td>
<td>203.65</td>
</tr>
<tr>
<td>q range, (Å(^{-1}))</td>
<td>0.0054</td>
<td>⎯</td>
<td>0.0054</td>
</tr>
<tr>
<td></td>
<td>0.1419</td>
<td>⎯</td>
<td>0.1401</td>
</tr>
<tr>
<td>GNOM total estimate</td>
<td>0.896</td>
<td>⎯</td>
<td>0.899</td>
</tr>
<tr>
<td>Volume of correlation ((V_c)) and (M_W) estimate</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Volume of correlation (V_c) (Å(^2))</td>
<td>⎯</td>
<td>1062</td>
<td>⎯</td>
</tr>
<tr>
<td>(M_W) estimate (kDa)*</td>
<td>159 ± 5</td>
<td>⎯</td>
<td>160 ± 5</td>
</tr>
<tr>
<td></td>
<td>α-Helix</td>
<td>β-Strand</td>
<td>β-Turn</td>
</tr>
<tr>
<td>--------</td>
<td>---------</td>
<td>----------</td>
<td>--------</td>
</tr>
<tr>
<td>ApoE2</td>
<td>57.3%</td>
<td>3.5%</td>
<td>14.1%</td>
</tr>
<tr>
<td>ApoE3</td>
<td>59.5%</td>
<td>3.3%</td>
<td>13.0%</td>
</tr>
<tr>
<td>ApoE4</td>
<td>58.4%</td>
<td>3.4%</td>
<td>13.5%</td>
</tr>
</tbody>
</table>

Table 3. Secondary structure analysis of recombinant ApoE isoforms from CD spectroscopy

a Calculated by CONTIN programme at DichroWeb (http://dichroweb.cryst.bbk.ac.uk/html/home.shtml)

The normalized root mean square deviation (NRMSD) is an estimate of the goodness-of-fit (NRMSD<0.1 is a necessary but not sufficient condition.

α-helical content was similar between the three ApoE isoforms (within experimental error).
<table>
<thead>
<tr>
<th>Tm (°C)</th>
<th>[Θ]_{222 \text{ nm at 37°C}} (deg.cm$^2$.dmol$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ApoE2</td>
<td>60.28 ± 1.18 -19181 ± 649</td>
</tr>
<tr>
<td>ApoE3</td>
<td>52.39 ± 1.92 -19809 ± 961</td>
</tr>
<tr>
<td>ApoE4</td>
<td>51.32 ± 4.32 -19092 ± 773</td>
</tr>
</tbody>
</table>

Table 4a. Thermal denaturation parameters from Boltzmann sigmoidal equation curve

Tm (°C) represents the temperature at which there was 50% change in α-helical content.

Non-linear best-fit parameters were determined on GraphPad Prism.

<table>
<thead>
<tr>
<th>Fraction unfolded at 37°C</th>
<th>%Fraction unfolded</th>
</tr>
</thead>
<tbody>
<tr>
<td>ApoE2</td>
<td>0.1 ± 0.07</td>
</tr>
<tr>
<td>ApoE3</td>
<td>0.13 ± 0.03</td>
</tr>
<tr>
<td>ApoE4</td>
<td>0.17 ± 0.06</td>
</tr>
</tbody>
</table>

Table 4b. Fraction unfolded of ApoE at 37°C

Fraction unfolded of ApoE at 37°C were calculated from the fraction unfolded curves given by equation 7 (Figure 3c). Non-linear best-fit parameters were determined using GraphPad Prism.
Table 4c. Chemical denaturation parameters from double sigmoidal curve

<table>
<thead>
<tr>
<th>Protein</th>
<th>[GuHCl]_{50, I} (M)</th>
<th>[GuHCl]_{50, U} (M)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ApoE2</td>
<td>0.69 ± 0.03</td>
<td>2.84 ± 0.04</td>
</tr>
<tr>
<td>ApoE3</td>
<td>0.71 ± 0.03</td>
<td>2.66 ± 0.07</td>
</tr>
<tr>
<td>ApoE4</td>
<td>0.58 ± 0.04</td>
<td>2.65 ± 0.08</td>
</tr>
</tbody>
</table>

[GuHCl]_{50, I} is the concentration needed for 50% of the protein to unfold to an intermediate state. [GuHCl]_{50, U} is concentration needed for 50% of the intermediate state to reach the unfolded state. Non-linear best-fit parameters were determined using GraphPad Prism.
A.

B.

C.

D.