Protein Misfolding and Degenerative Diseases



Aaron Trowbridge

Literature Review 27th January 2021



Global mortality linked to disease



Note: Shown is period life expectancy at birth, the average number of years a newborn would live if the pattern of mortality in the given year were to stay the same throughout its life.

The major causes of global deaths have changed profoundly over the last 200 years

Global mortality linked to disease



Bacterial infections (e.g. tuberculosis)



Viral infections (e.g. influenza)



Proteopathic diseases (e.g. AD, Parkinson's)

Approximately **40%** of human diseases may be attributable to protein misfolding

Global mortality linked to disease



Proteins participate in virtually every process within the cell

Functional impairment of proteins can be devastating

What are the genetic and molecular causes for incorrectly formed proteins?



Protein Misfolding and Degenerative Diseases

- Protein function and three-dimensional structure
- The energetic funnel
- Chaperones
- Misfolding in neurological diseases
- Misfolding in other diseases
- Looking forward and treatments
- DeepMind



In 1917, the German chemist Hermann Staudinger proposed that organic molecules such as proteins were organised into polymers, for which he received the Nobel prize in 1953.









Protein folding is driven by a series of interactions

Protein folding is fast...but should it be?



Levinthal's paradox

A protein with *N* amino acids should have 10^{2N} degrees of freedom. Thus a small protein (e.g. 150 amino acids) would have 10³⁰⁰ degrees of freedom. There is not enough time in the universe to try each of these combinations, and yet the protein is **folded within a second**.



This surface represents the 'dry state' where protein folding is slow and challenging

Protein folding is governed by many factors including **pH**, **ionic strength**, **temperature**, and **cofactors**



The rearrangement of water around unfolded protein chains releases energy



- Protein folding is mainly driven by the 'hydrophobic effect' through the burial of non-polar side chains
- Water adjacent to a hydrophobic group loses a hydrogen bond, therefore has increased enthalpy
- To overcome this the network expands to form lower-density water with a lower entropy

...an alternative view

Protein conformation tends to minimise the disruption of the water matrix



- Collapse of water around hydrophobic residues is accompanied by structural
 H-bond formation and the formation of strong ion-paired salt links
- A driving force for this is the release of water to be available for the hydration of other solutes, hence **maximising entropy**

Fernández, A., The principle of minimal epistemic distortion of the water matrix and its steering role in protein folding, *J. Chem. Phys.*, **139**, 085101, (2013); Fernández, A., Kardos, J., Goto, Y., Protein folding: could hydrophobic collapse be coupled with hydrogen-bond formation? *FEBS Lett.*, **536**, 187–192, (2003); Chen. S., Itoh, Y., Masada, T., Shimizu, S., Zhao, J., Ma, J., Nakamura, S., Okuro, H., Noguchi, H., Uosaki, K., Aida, T., Subnanoscale hydrophobic modulation of salt bridges in aqueous media, *Science*, **348**, 555–559, (2015); Harano, Y., Roth, R., Kinoshita, M., On the energetic of protein folding in aqueous solution, *Chem. Phys. Lett.*, **432**, 275–280, (2006)



Potential Energy Funnel

- When hydrated, the potential energy landscape is considerably smoother
- Minimum energy conformation can be attained rapidly
- Potential energy barriers lowered due to the ease with which water molecules can lubricate the movement of the amino acid backbone
- Amino acid side chains reduce the hydration of peptide groups via shielding, promoting folding

...a case study

Case study:

How do spiders avoid premature aggregation of silk (spidroin)?



Euprosthenops australis



Spidroins are stored at remarkably high concentrations (30–50% w/w) in the spider silk gland

Spidroin is amphiphilic

Most of the charged residues are exposed in the dimer; **basic** residues at the positive pole and **acidic** residues at the negative poles



Spidroin is amphiphilic

Clustering of acidic residues increases the pK_a of the side chain residues by several pH units



N terminal domain (NT)

C terminal domain (CT)

Repetitive region

Spidroins form **micelles**, where the hydrophilic NT and CT domains shield hydrophobic regions





This allows the intrinsic pH gradient of spider silk glands to regulate silk formation



pH = 7 self-assembly is slow (~2 h)

pH = 6 self-assembly is rapid (<5 min)





Translation is slow (100aa ~ 25 s) compared to folding and so occurs co-translationally

Protein folding starts in the ribosome tunnel



Wilson, D. N., Beckmann, R., The ribosomal tunnel as a functional environment for nascent polypeptide folding and translational stalling, Curr. Opin. Struct. Biol. 21, 274–282, (2011)



Uneven tunnel geometry and electrostatics speed up and slow down folding

Protein folding does not occur by one path



Hen lysozyme (14.3 kDa)

As proteins increase in size, the barriers to reorganisation in the collapsed state are **larger** - resulting in multiple 'pathways' to the native state

Real-life kinetics differ significantly from single exponential behaviour, and distinct folding populations are created by the accumulation of longlived intermediates along particular pathways





Significantly populated intermediates that have both native and non-native interactions

Resulting in extreme heterogeneity

70% molecules populate an intermediate with persistent structure in the *α*-domain,
followed by folding to the native state within 400 ms



20% molecules fold to the native state <100 ms 10% molecules fold extremely slowly

(limited by proline isomerization)



Yellow Trajectory:

The α and β domains form concurrently, and the free energy decreases almost monotonically towards the native state (the 'fast track')

Red Trajectory:

Formation of the α domain precedes the β domain, and becomes trapped in a minimum. The α domain partially unfolds (reversal in the trajectory) to form the β domain.

Bifurcating pathways can lead to proteins becoming stuck in energetically minimal structures that aren't the native conformation





Specialized proteins called molecular chaperones help complicated or unstable proteins find their native conformation



Molecular chaperones were first mentioned in 1978 by Ron
 Laskey, who found that nucleoplasmin (a protein found in the nucleus of the cell) is able to bind to histones



 Laskey observed that nucleoplasmin acted like a chaperone, preventing the aggregation of folded histone proteins in DNA during nucleosome assembly HSPs regulate targeting and degradation of misfiled proteins at the lumen of the ER



Some proteins can evade the ER 'quality control system'

As millions of copies of each protein are made during our lifetimes, sometimes a random event occurs leading to a misfolded protein (also called toxic conformations)





When a protein becomes toxic, an extensive conformational change occurs and it acquires a motif known as the beta sheet



The conformational transition from alpha helix to beta sheet exposes hydrophobic amino acid residues and promotes protein aggregation



β-amyloid fibrils



The misfolded proteins attach to other healthy proteins, building a template that rapidly grows, resembling a crystallisation process

This continual cycle occurs until the cell dies and disperses the infectious protein to surrounding cells, which ultimately leads to damage





1997 Nobel prize for medicine The term **prion** (**pr**oteinaceous infect**ion**) was coined by Stanley Prusiner in 1982 to explain small mutated proteins responsible for various unusual brain diseases. He was not believed for nearly two decades.



Scrapie is a fatal degenerative disease found in sheep and goats. The name derives from the compulsive scraping off of the animals fleeces.

There is no cure

The cause of scrapie is unknown and a matter of debate

The agent is very difficult to destroy with heat, radiation and disinfectants, it can remain in the soil for >10 years, does not evoke any detectable immune response, and has a long incubation period of between 18 months and 5 years




Prions are also responsible for BSE (mad cow disease), whose infective form can cause Creutzfeldt-Jakob disease in humans; and **kuru**, the only epidemic human prion disease known





Lasmézas, C. I., Fournier, J.-G., Nouvel, V., Boe, H., Marcé, D., Lamoury, F., Kopp, N., Hauw, J.-J., Ironside, J., Bruce, M., Dormont, D., Deslys, J.-P., Adaptation of the bovine spongiform encephalopathy agent to primates and comparison with Creutzfeldt-Jakob disease: implications for human health, *Proc. Nat. Acad. Sci.* **98**, 4142–4247, (2001)

Kuru was discovered in the late 1950's among populations of the Fore tribe of the eastern highland of Papua New Guinea

Kuru was transmitted by funerary cannibalism, where the deceased were traditionally cooked and eaten.

Women and children, who usually ate the brain, were the most severely effected.





In 2009, a naturally occurring variant of the prion protein was found (in areas where Kuru was most widespread) that conferred strong resistance

Heterozygous polymorphism in *PRNP* (coding either M or V) is protective against the development of CJD and Kuru. (All known cases of variant CJD have been MM)

sCJD = natural (most common); vCJD = consuming meat (rare); iCJD = medical contamination (rare)

> Mead, S., Whitfield, J., Poulter, M., Shah, P., Uphill, J., Campbell, T., Al-Dujaily, H., Hummerich, H., Beck, J., Mein, C. A., Verzilli, C., Whittaker, J., Alpers, M. P., Collinge, J., A novel protective prion protein variant that colocalizes with Kuru exposure, *N. Engl. J. Med.* **361**, 2056–2065, (2009)

What does the prion protein (PrP) do?



PrP^c is a 208 residue GPI-anchored protein present on the cell surface

The exact biological function of PrP^c is unknown. It has been implicated in signal transduction, Cu-binding, synaptic transmission, and induction of apoptosis, but is <u>not essential</u>

The exact composition of the protein form causing the prion diseases remains to be fully understood

Post-mortem studies show a poor correlation between the load of amyloid like deposits in the brain and the severity of clinical symptoms

Alternative forms of PrP have been reported to have important roles in prion-mediated neurodegeneration



Prion disease

Several lines of evidence also implicate misfolded oligomers as playing a role in prion disease

Protein aggregates have different roles in distinct diseases



Alzheimer's

Mode of transmission: sporadic (95%), inherited (5%) Clinical features: progressive dementia Affected regions: hippocampus, cerebral cortex Proteins involved: amyloid-β, tau proteins Cellular location: extracellular, cytoplasmic

Extracellular amyloid plaques (white arrows) deposited around the cerebral vessel walls are comprised of Aβ. Intracytoplasmic neurofibrillary tangles (yellow arrows) are comprised of hyperphosphorylated tau protein.

> Glenner, G. G., Wong, C. W., Alzheimer's disease: initial report of the purification and characterisation of a novel cerebrovascular amyloid protein, *Biochem. Biophys. Res. Common.* **120**, 885–890, (1984); Grundke-Iqbal, I., Iqbal, K., Quinlan, M., Tung, Y. C., Zaidi, M. S., Wisniewski, H. M., Microtubule-associated protein tau. A component of Alzheimer paired helical filaments. *J. Biol. Chem.* **261**, 6084–6089, (1986)



Parkinson's

Mode of transmission: Mostly sporadic Clinical features: movement disorder Affected regions: substantia nigra, hypothalamus Proteins involved: α-synuclein Cellular location: cytoplasmic

Cytoplasm of neurons from the substantia nigra contain aggregates called Lewy bodies (white arrows). The major constituent of these aggregates are fragments of a protein named α-synuclein.



Huntington's

Mode of transmission: Inherited (autosomal dominant) Clinical features: dementia, motor, and psychiatric Affected regions: striatum, cerebral cortex Proteins involved: huntingtin Cellular location: nuclear

Intranuclear deposits of a polyglutamine-rich version of huntingtin protein (white arrows) are a typical feature of brains from patients with HD.



Amyotrophic lateral sclerosis (ALS)

Mode of transmission: sporadic (90%), inherited (10%) Clinical features: movement disorder Affected regions: motor cortex, brainstem Proteins involved: superoxide dismutase Cellular location: cytoplasmic

In ALS, patients have aggregates, mainly composed of superoxide dismutase (SOD1), in the cell bodies and axons of motor neurons.

Bruin, L. I., Houseweart, M. K., Kato, S., Anderson, K. L., Anderson, S. D., Ohama, E., Reaume, A. G., Scott, R. W., Cleveland, D. W., Aggregation and motor neuron toxicity of an ALSlinked SOD1 mutant independent from wild-type SOD1, *Science*, **281**, 1851–1854, (1998) No other protein conformational disorders (aside from prion disease) have been convincingly shown to be transmissible*

*Some animal models have shown that AD pathology can be accelerated by the injection of enriched protein homogenate:

What causes misfolding amongst other proteins?



Internal hydrophobic region (AAs 17–21; LVFF) is implicated in the early steps of Aβ misfolding, with aggregation at the C-terminus

α-Synuclein fibrillogenesis, although less-well studied, stems from the central hydrophobic region (AAs 61–95)



In Huntington's and ALS, both disease and protein aggregation are associated with an inherited expansion of CAG (glutamine) codon repeats Trinucleotide repeat disorders are genetic disorders caused by mutation in which repeats of three nucleotides increase in copy number until they become unstable



Trinucleotide repeat disorders generally show genetic anticipation: their severity increases with each successive generation that inherits them

DNA slippage occurs during DNA synthesis due to misalignment in hybridisation



Aggregation of huntingtin in vitro depends on the length of the polyglutamine repeat



Polar zipper model

β-sheets are formed and stabilized by
the collective strength of cooperative
hydrogen bonding involving the amide
group of the glutamine residue

What is the mechanism of neuronal death induced by protein misfolding?



Loss-of-function hypothesis





Amyotrophic lateral sclerosis (ALS)

In ALS, the protein (that becomes misfolded) SOD1 catalyses the conversion of superoxide to hydrogen peroxide

Depletion of SOD1 by misfolding and aggregation could lead to accumulation of superoxide radicals



 SOD1 knockout mice show no degeneration of motor neurons or oxidative damage

SOD1 mutants with different activity do not correlate to age of onset or severity of disease

Reaume, A. G., Elliott, J. L., Hoffman, E. K., Kowall, N. W., Ferrante, R. J., Siwek, D. R., Wilcox, H. M., Flood, D. G., Beal, M. F., Brown Jr., R. H., Scott, R. W., Snider, W. D., Motor neurons in Cu/ Zn superoxide dismutase-deficient mice develop normally but exhibit enhanced cell death after axonal injury, *Nat. Genet.*, **13**, 43–47, (1996); Borchelt, D. R., Guarnieri , M., Wong, P. C., Lee, M. K., Slunt, H. S., Xu, Z.-S., Sisodia, S. S., Price, D. L., Cleveland, D. W., Superoxide dismutase 1 subunits with mutations linked to familial amyotrophic lateral sclerosis do not affect wild-type subunit function, *J. Biol. Chem.* **270**, 3234–3238, (1995)

Gain-of-toxic-activity hypothesis



Percentage of cell deaths induced by 48-h-aged HypF-N aggregates (solid) and control protein (grey)



β-sheet oligomerization of non disease-related proteins is cytotoxic, indicating that misfolding and aggregation of any protein results in inherent toxicity

Bucciantini, M., Giannoni, E., Chiti, F., Baroni, F., Formigli, L., Zurdo, J., Taddei, N., Ramponi, G., Dobson, C. M., Steffani, M., Inherent toxicity of aggregates implies a common mechanism for protein misfolding diseases, *Nature* **416**, 507–511, (2002)

Gain-of-toxic-activity hypothesis



Activation of a signalling pathway



Extracellular aggregates might activate a signal transduction pathway that leads to apoptosis by interacting with specific cellular receptors

RAGE (receptor for advanced glycation end products) is a transmembrane receptor of the immunoglobulin superfamily

- Upon binding of a ligand (Aβ), RAGE signals activation of NF-κB9, which causes an inflammatory response
- In AD, where lots of Aβ is present, this generates a positive feedback loop leading to chronic inflammation and cell death



Recruitment of cellular factors



Intracellular aggregates might damage cells by recruiting factors that are essential for cell viability into the fibrillar aggregates



Distribution of the proteasome in neurons of the nucleus pontis centralis from an SCA1 patient (L) and control (R). Note the redistribution of the proteasome to the sites of ataxin-1 aggregation in the patient. Components of the proteasome,

chaperone proteins, cytoskeletal proteins and transcription factors have been found in huntingtin and α-synuclein aggregates

Cummings, C. J., Mancini, M. A., Antalffy, B., DeFranco, D. B., Orr, H. T., Zoghbi, H. Y., Chaperone suppression of aggregation and altered sub cellular proteasome localisation imply protein misfolding in SCA1, *Nat. Gen.* **19**, 148–154, (1998); Li, K., Ito, H., Tanaka, K., Hirano, A., Immunocytochemical co-localization of the proteasome in ubiquitinated structures in neurodegenerative diseases and the elderly, *J. Neuropathol. Exp. Neurol.* **56**, 125–131, (1997)



Formation of

Membrane disruption and depolarisation mediated by ion-channel formation, resulting in alteration of ion homeostasis and cell death



At zero membrane potential, a net negative current corresponding to the movement of K+ is shown

- Synthetic AβP forms cation-selective channels across planar lipid bilayers
- Mediation of cell death via discharge of cellular membrane potential

Arispe, N., Rojas, E., Pollard, H. B., Alzheimer disease amyloid β-protein forms calcium channels in bilayer membranes: Blockade by tromethamine and aluminium, *Proc. Natl. Acad. Sci.* **90**, 567–571, (1993); Lin M. C., Mirzabekov, T., Kagan B. L., Channel formation by neurotoxic prion protein fragment, *J Biol. Chem.*, **272**, 44–47, (1997)

Oxidative stress



The production of free radical species results in protein and lipid oxidation, elevating levels of intracellular calcium and mitochondrial disfunction



Primary cells exposed to Aβ for 23h (L) and control (R). Staining with peroxide active dye.

- Aβ causes increased levels of hydrogen peroxide and lipid peroxides in cells
- Breakdown of peroxide into hydroxyl radical results in cellular necrosis

Behl, C., Davis, J. B., Lesley, R., Schubert, D., Hydrogen peroxide mediates amyloid β-toxicity, *Cell*, **77**, 817–827, (1994); Hsu, L. J., Sagara, Y., Arroyo, A., Rockenstein, E., Sisk, A., Mallory, M., Wong, J., Takenouchi, T., Hashimoto, M., Masliah, E., α-synuclein promotes mitochondrial deficit and oxidative stress, *Am. J. Pathol.*, **157**, 401–410, (2000); *For review, see:* Lin, M. T., Flint Beal, M., Mitochondrial dysfunction and oxidative stress in neurodegenerative diseases, *Nature*, **443**, 787–795, (2006)

The hypothesis that aggregates are toxic has been challenged by recent studies

In some animal models of AD, TSE, HD and ataxias, cerebral damage and clinical symptoms have been detected before protein aggregates

Neuropathological analysis of AD and PD patients show Lewy bodies and tangles that are healthier than neighbouring cells

Amyloid-like plaques and Lewy bodies are

found in people without evident neuronal loss or clinical signs of AD or PD

The process of misfolding and **early stages** of oligomerization, rather than deposition of aggregates, may be the real culprits in neurodegeneration

Hartley, D. M., Walsh, D. M., Ye, C. P., Diehl, T., Vasquez, S., Vassilev, P. M., Teplow, D. B., Selkoe, D. J. Protofibrillar intermediates of amyloid beta-protein induce acute electrophysiological changes and progressive neurotoxicity in cortical neurons, *J. Neurosci.* **19**, 8876–8884, (1999); Goldberg, M. S., Lansbury, P. T. Jr., Is there a cause-and-effect relationship between α-synuclein fibrillation and Parkinson's disease, *Nat. Cell Biol.* **2**, E115–E119, (2000)

Brain inflammation hypothesis



Abnormal protein aggregates act as irritants and cause a chronic inflammatory reaction in the brain that leads to neuronal death and synaptic changes



HD cerebral caudate (L) and control (R) showing enlarged microglia and enhanced immnoreactivity

Recent studies show both positive and negative effects of inflammation

What strategies exist to combat conformational diseases?

Despite impressive progress in understanding the pathogenesis of neurodegenerative diseases, **none** of these disorders can be successfully treated

At least four approaches have been proposed to attack protein misfolding and aggregation



Stabilisers of protein folding

Chemical chaperones known to stabilise native protein conformations



Tafamidis *Vyndaquel (Pfizer)* Received FDA approval in 2019 for the treatment of cardiomyopathy via the stabilisation of the protein transthyretin

At least four approaches have been proposed to attack protein misfolding and aggregation

Tafamidis stabilises the β-sheet assembly in tetrameric species





At least four approaches have been proposed to attack protein misfolding and aggregation



β-sheet breakers

Short synthetic peptides

designed to arrest the folding of the polypeptide in the β -sheet

Typically unstable in vivo (therapeutic limitations)

At least four approaches have been proposed to attack protein misfolding and aggregation



Competitive inhibitors (monomers and oligomers)

Compounds that block the interaction between monomers and molecules that tack at the edge of β-sheet aggregates



Fabry disease is a rare genetic disorder caused by mutations to the enzyme α-GalA , leading to **misfolding**

Migalastat *Galafold (Amicus)*

Binding of migalastat to α-GalA shifts the folding behaviour towards the proper conformation, resulting in a functional enzyme

At least four approaches have been proposed to attack protein misfolding and aggregation

Clearance enhancers

Aggregates of synthetic misfolded protein can be used as antigens to induce immune response, producing antibodies to clear them



At least four approaches have been proposed to attack protein misfolding and aggregation



Prof. Gal Bitan (UCLA)

"CLR01 acts by decreasing synuclein in the brain by wrapping around lysine chains and remodelling the assembly into non-toxic structures"
Protein Function and Three-Dimensional Structure

Alzheimer's disease therapy: a critical appraisal

Protein Function and Three-Dimensional Structure



Protein Function and Three-Dimensional Structure

Alzheimer's disease therapy is based on the amyloid cascade hypothesis

Amyloid cascade hypothesis



Amyloid precursor protein

(APP) is cleaved by β secretase 1 (BACE1)
 generating a soluble
 extracellular fragment (sAPPβ)
 and membrane bound
 fragment (C99)

C99 is cleaved by an

enzymatic complex of four proteins, collectively termed **γsecretase** releasing Aβ and an intracellular peptide known as amyloid intracellular domain (AICD)

Aβ aggregates to
 ultimately form plaques
 (hallmarks of AD)

 Intracellular deposition of hyperphosphorylated tau

protein in neurofibrillary tangles (NFTs) leads to progressive cytoskeletal changes and disrupts axonal transport

Karran, E., Mercken, M., De Strooper, B., The amyloid cascade hypothesis for Alzheimer's disease: an appraisal for the development of therapeutics. Nat. Review Drug Discov. 10, 698–712 (2011)

Amyloid cascade hypothesis

Many opportunities for therapeutics?

Putative disease modifying therapies



Putative disease modifying therapies



Almost all drugs that decrease the production of A β or increase A β brain clearance have failed to improve cognitive outcomes despite reducing brain A β

Putative disease modifying therapies

Is it time to rethink the amyloid cascade hypothesis?

The sepsis parallel

Sepsis is a complex disorder that develops as a dysregulated host response to an infection and is associated with high risk of death



Leukocytosis is typical **finding**

The sepsis parallel

Sepsis is a complex disorder that develops as a dysregulated host response to an infection and is associated with high risk of death



Leukocytosis is typical **finding**

Treatment is aimed at **underlying infection**

The sepsis parallel



Aβ accumulation might be a reaction of the brain to **neuronal damage** (e.g. *mitochondrial dysregulation*)

An alternative option for AD therapeutics is to address the modifiable risk factors for developing AD



Neurons are dependent on the

neurotrophic properties of insulin

Resistance to the effects of insulin make neurons vulnerable to stress **resulting in impairments in function**

In the aged brain, insulin might **lose its effectiveness** as a growth factor. Therefore, **anti-diabetes drugs** might be an effective approach for treating AD

Several studies have identified a link between the brain's immune system and AD, of which **microglia** have a central role

Microglia might be activated by peripheral inflammation and/or **gut microbiota**



The future of gene modifying therapies

Autosomal dominant mutations were found in the genes coding for APP, PSEN1, and PSEN2. Transgenic mice that expressed these mutations developed brain Aβ plaques and memory deficits (1990s)

LETTERS TO NATURE								
LETTERS TO NATURE					REPORTS			REPORTS
					KLIOKIS			ALL ON LO
sequence of 17 amino acids within H5 (432-448) as a β hairpin that extends = 20 Å into the membrane. The interaction of	channel compl	ex is a tetramer, changing a single residue would	mean onset age for the family, a combined standard	dependent reduced penetrance, and thus cannot be	American Physicians Fellowship. We thank D. Noch-		the absence of TCRa, the number of	that corresponded to amino acids 88 to 101 of anti-CD3 ₆ (14), or anti-ζ (mixture of mAbs 528 and 551) (6) that had been set to a set the set of the set
β -hairpins from each subunit of the channel could form the	of threonine t	o serine, for example, could increase the local	years), and an utimate penetrance of 100%. For the	 Z7. XD. Cai, T. E. Golde, S. G. Younkin, Science 259 	in, w. sum, w. bait, and G. Germer for neuropatho- logic characterization of many of the families, E. Loo-		CD4+CD8+ thymocytes is normal (10,	(5); antibody 551 was generated to a peptide that to protein A agarose beads (Sigma). Immuno-
central pore as a β barrel that spans the distance between the	pore diameter	by the equivalent of two methyl groups. Thus,	mined by an age-dependent penetrance model with	 514 (1963); M. Cilton et al., Nature 380, 672 (1982) M. Cilton et al., Proc. Natl. Acad. Sci. U.S.A. 91. 	mis and H. Utsugi for technical help, H. Lipe and T. Lampe for obtaining blood samples and medical		CD4+CD8+ thymocytes were reported for	and recognizes only the ζ protein (5). for 5 min in SDS sample buffer containing 2-mer-
been proposed for some less selective ion channels such as the	if the altered r	esidue contributes to a limiting constriction. The	a phenocopy correction that assumes a cumulative normal age-of-onset function with a mean of 110	 11993 (1994); N. Suzuki et al., Science 264, 1336 (1994); A. Tamaoka et al., J. Biol. Chem. 269, 32721 	records, D. Breiter for assistance with computations, and M. Boehnke and A. Motulsky for reading the		TCR $\beta^{-/-}$ mice (10), which suggests that	7. P. E. LOVE, E. W. Snores, H. Westphal, A. Singer, unpublished data. singer, sional SDS-polyacrylamide gel electrophoresis
voltage-dependent anion channel and porins ^{18,19} .	absence of cha	nges in K ⁺ conductance in our mutants suggests	years and a standard deviation of 14 years (12). Possible heterogeneity was examined with the ad-	 (1994). 28. D. Games et al., Nature 373, 523 (1995). 	manuscript prior to publication. Informed consent was obtained from each subject or next of kin with		into the CD4-CD8 developmental pathway	 D. J. Paterson and A. F. Williams, J. Exp. Med. (PAGE) (13%). Gels were fluorographed with 166, 1603 (1987); H. R. MacDonald, R. C. Budd, dimethyl sulfoxide/2,5-diphenyloxazole, dried,
and NH ⁴ , without removing the ability of the channel to exclude	by the mutati	ons, remains limiting for K* flux but has less	mixture test [S. E. Hodge, C. E. Anderson, K. Neiswanger, R. S. Sparkes, D. L. Rimoin, ibid. 35,	 Supported by grants from the National Institute or Aging for the Azheimer's Disease Research Center 	approval of the University of Washington Human Subjects Review Committee, We also thank the		or, alternatively, expansion of the pool of	R. C. Howe, Eur. J. Immunol. 18, 519 (1988); K. Shortman, A. Wilson, M. Egerton, M. Pearse, R. 18, D. R. Greaves, F. D. Wilson, G. Lang, D. Kioussis.
Na ⁺ . These data provide evidence that different structural mechanisms are involved in selecting against cations that are	influence on R	b ⁺ and NH ₄ ⁺ , which have lower charge densities.	1139 (1983)]. 18. Unites otherwise noted, published allele frequencies	at the University of Washington (AG05136, G. M. Martin, director), the National Center for Human Ge-	many family members who generously cooperated in this study, the American Historical Society of Ger-		pression of both β and ζ subunits.	Scollay, Cell. Immunol. 113, 462 (1988); P. Hugo, G. A. Waanders, R. Scollay, K. Shortman, R. L. 19. For multicolor flow cytometry (FCM), thymocytes or
larger or smaller than K ⁺ . These results are consistent with	of Na ⁺ and C	a ²⁺ channels may reveal the molecular bases of	are from the Genome Database. 19. J. Ott. Am. J. Hum. Genet. 51, 283 (1992).	nome Research (HG00835), Veterans Atfairs Re search Funds, an National Research Studies Alz-	mans from Russia, and T. Kloberdanz, G. Miner, and R. Cook-Deegan.		We also did not expect to find SP T cells	Boyd, Int. Immunol. 2, 209 (1990). 9. H. Kishi et al., EMBO J. 10, 93 (1991): E. W. Immunol. 2, 209 (1990). Immunol. 2
previous models in which multiple barriers exploit both steric bindrance and differences in ionic charge density ^{2,20} If the	the energy bar selectivity	riers that limit permeation and determine ionic	 Marker order for some loci was from G. Gyapay et al., Not. Count. 7, 246 (1004). This order with continues. 	heimer's post-doctoral fellowship (F32 AG05635) the American Health Assistance Foundation and the	2. kms 1995: accented 14. kit/1995		These cells were not vδ T cells: they did	Shores et al., J. Immunol. 150, 1263 (1993). 10. P. Mombaerts et al. Nature 560, 225 (1992) Children of contracting the End the second
			(sex-averaged) distances is D1S28-4-D1S422-4- D1S412-2 D1S422-4-				not express surface TCR8 (7). The genera-	 K. L. Philpott et al., Science 256, 1448 (1992). FCM, cells were sequentially incubated with fluo- recently incubated with fluo- incubated with fluo- incubated with fluo- incubated with fluo
Received 3 December 1990; accepted 2 January 1991.	15. Taylor, P. S. J. Phy	Hot, Long, 389, 437-447 (1967).	D13412-1-D13308-6-D13249-1-D13249-1- D13205-2-D13425/D13217-6-D13229-1-	Candidata Cono for the	Chromosomo 1 Familial		tion of SP T cells is thought to require positively selecting TCR signals that are	 R. T. Kubo, W. Born, J. W. Kappler, P. Marrack, M. Picera, J. Improv. 142, 2232 (1980). B. T. Kubo, W. Born, J. W. Kappler, P. Marrack, M. Biodels, then biotinylated antibodies, then Texas
 Bezanila, F. & Armstrong, C. M. J. gen. Physiol. 60, 558–628 (1972). Hile, B. & Schwarz, W. J. gen. Physiol. 72, 409–442 (1978). 	17. Hoshi, T., Zagotta, 18. Blachly-Dyson, E., I	W. & A Molet, R. W. Solence 250, 533-538 (1990). erg. S. Colombri M. & Forte, N. Solence 251, 2233-1236 (1990).	2-D15459-2-D15446-4-D15235. As part of this		Disease Leave		induced in CD4+CD8+ thymocytes that	14. O. Leo, M. Foo, D. Sachs, L. E. Samelson, J. A. Buschene Pere Med Aced Science and Control of the Control o
 Amarong, C. M. Q. Wei, Begryn, Y. 179-210 (1975). Latore, R. & Miler, C. J. memb. Biol 71, 11-30 (1983). Markines, B. & Miler, C. G. Science, 200, 2720 (1997). 	19. Van der Ley, 20. Eisenman, G.		the complete CEPH mapping panel (40 tamilies) and the	Aizheimers			are TCR ^m and CD5 ^m . The CD4 ⁺ CD8 ⁺ thymocytes in <i>I</i> ^{-/-} mice are CD5 ^{lo} and	(1987). were collected on 5 × 10 ⁴ to 2.5 × 10 ⁵ viable cells
 Matchina C, K. & Koter, G. Schwarz, T. L., Jan, Y. N. & Jan, L. Y. Science 237, 770–775 (1987). Tempel, B. L., Papazian, D. M., Schwarz, T. L., Jan, Y. N. & Jan, L. Y. Science 237, 770–775 (1987). Schwarz, T. L. Terroel, R. L., Papazio, D. M. Jan, Y. N. & Jan, L. Y. Nature 381, 137, 142 (1988). 	21. Tempel, B. L. 22. Stühmer, W.	ARTICLES	computer program CHIMAP [E. R. Lander and P. Green, Proc. Natl. Acad. Sci. U.S.A. 84, 2363, (1987);	Ephrat Levy-Lahad,* Wilr		LETT	express only barely detectable quantities of	 Soumern and Normern bioting and hybridizations were performed essentially as described [F. M. as determined by electronic gating on torward light scatter and propidium iodide exclusion and ana-
 Timpe, L. C. <i>et al.</i> Nature 333, 143–145 (1988). Timpe, L. C., Jan, Y. N. & Jan, L. Y. Neuron 1, 659–667 (1988). 	23. Butler, A., W 24. Frech, G. C.,		the order was D1S229-5.7-D1S320-2.6-D1S213-1- D1S479/D1S439-5.4-D1S103.	Donna M. Romano, Junko			surface TCR; it was therefore expected that	Ausuber et al., Eds., Short Protocols in Molecular Biology (Wiley, New York, ed. 2, 1992), pp. 2.24- Division of Computer Research and Technology.
 Nerson, L. E., Tanouye, M. A., Davidson, N. & Rudy, B. Proc. natr. Acad. Sci. U.S.A. 86, 8723–5727 (1988). 	(1989). 25. Kunkel, T. A.		 Initial power analysis of the families (which did not include the neuropathologic samples) indicated that 	Chang-en Yu, Paul D. Jondr		Initial studies of the \$782	CD4 ⁺ CD8 ⁺ stage of differentiation. We do	2.30 and 4.23–4.24). NIH. The mAbs used for FCM analysis included 16. P. Fort et al., Nucleic Acids Res. 13, 1431 (1985). FITC-anti-Leu4 and biotin-anti-Leu4 (BDIS):
 Zagotta, W. N., Hoshi, T. & Aldrich, R. W. Proc. natn. Acad. Sci. U.S.A. 86, 7243–7247 (1989). Guy, H. R. in Monovalent Cations in Biological Systems (ed Pastemae, C. A.) (CRC Press, Cleveland. 	ACKNOWLEDGEM	Cloning of a gene hearing missoned	with a reasonably polymorphic marker (PIC = 0.7; D. Botstein, R. L. White, M. Skolnick, R. W. Davies, Am	David Galas Ellen Nemens	Familial Alzheimer's disease	genes with similar sequences ¹ . Searc	not yet know whether the unusual periph-	 Thymocytes (2.5 × 10²) were metabolically la- beled for 30 min in methonine-free RPMI (Sig- CD8 (53-6.72: BDIS). FITC-anti-CD4 (Hm4-5; Pharmingen), biotin-anti- CD8 (53-6.72: BDIS). FITC-anti-CD4 (Hm4-5; Pharmingen), biotin-anti- CD8 (53-6.72: BDIS). FITC-anti-CD4 (Hm4-5; Pharmingen), biotin-anti- CD8 (53-6.72: BDIS).
Ohio, 1989). 13. Jan, L. Y. & Jan, Y. N. Nature 345, 672 (1990).	discussions and Fellowship to T.L.	eroning of a Solic scaling illiggelige	J. Hum. Gonet. 32, 314 (1980)], at # = 0.05, the mean Z = 2 01 with a maximum acceleration	Gerard D. Scheller	in kindreds with missense	database by using the BLASTX alig	erai SP 1 cells detected are derived from the SP cells in the thymus or whether they	ma) plus 10% fetal calf serum with [²⁶ S]methio- nine (1 mCl/ml) [J, S, Bonifacino et al., Nature H12: BDISty CDS (95,7 % DDS) (13); and Thy-1.2 (30- H12: BDISty CDS (95,7 % DDS)
24. Hene, B. J. gent Physiol 61, 669-686 (1973).	Medical Institute	mutations in early-onset familial	score of 7.55, and at 9 = 0.1, these lod scores would be 1.8 area 6.25 corestration for the formation of the		mutations in a gene on	subsequently R14600 and R05907)	are generated extrathymically by a cryptic	344, 247 (1990)). Labeled cells were solubilized at 4°C with buffer containing 0 15 M NaCI 50 mM
Commission of a mission	onest fami	······································	somewhat less than the 4.40 actually obtained, but	A candidate gene for the chromosome (STM2). The predicted aming acid server	abremeseme 1 related	to two different segments within the	developmental pathway. However, their pormal CD4/CD8 ratio is most consistent	tris-HCI (pH 7.4), leupeptin (20 μg/m), aprotinin (40 μg/m), deproteining the J1 ES cell line. We thank Y, Taka-
Segregation of a missense	poses for g	Alzheimer's disease	as the power analysis did not assume the availability of the neuropathologic samples, such a result was	recently cloned chromosome 14 AD ger	chromosome 1 related	nucleotide primers were designed fro	with their being of thymic origin.	NP-40. Labeled TCR proteins were immunopre-
mutation in the amyloid	AD and ma family with		expected. 22. Control allele frequencies for D1S479 were deter-	the substitution of an isoleucine for an	to the Alzheimer's	enerate probes to screen compleme	In conclusion, ζ performs a previously	ogenericul with anterioritiα (12), anterioritig (13), 1 June 1993; accepted 21 July 1993
precursor protein gene with	in this kind	R. Sherrington , E. I. Rogaev , Y. Liang , E. A. Rogaeva , G. Lev	mined by allele counting. Frequencies for grandpar- ents of members of pedigrees from the Centre	residue that is conserved in human S182	disease type 3 gene	size from 1 to 2.3 kilobases (kb). T	moting the generation or expansion of	
familial Alzheimer's disease	there is a p	M. Ikeda , H. Chi , C. Lin , G. Li , K. Holman , T. Tsuda , L. Mar	d'Etude du Polymorphisme Humain (CEPH) (244 control chromosomes), all the VG subjects (affect-	of missense mutations in AD subjects in		these clones confirmed that all were of the transcript (designated FS-D). The	CD4+CD8+ thymocytes and is critical for	Gene Dose of Apolipoprotein E Type 4 Allele
Taminal Alzhenner S uisease	minus of th	JF. Foncin [®] , A. C. Bruni [®] , M. P. Montesi [®] , S. Sorbi [®] , I. Rainero [*]	eds, at-risk, and spouses; 254 chromosomes), af- fected VG subjects (74 chromosomes), and VG	hypothesis that mutations in both are p	E. I. Rogaev*, R. Sherrington*, E. A. Rogaeva*,	transcript was mapped to human chro	prisingly, some SP T cells can be generated	and the Risk of Alzheimer's Disease in
Alison Goate*, Marie-Christine Chartier-Harlin*,	AD reveale This sugges	L. Nee", I. Chumakov , D. Pollen', A. Brookes', P. Sanseau",	spouses (32 chromosomes) are respectively: 102 bp. 0.013, 0.012, 0.0, and 0.063; 104 bp. 0.029		C. Lin*. K. Holman*. T. Tsuda*. L. Mar†. S. Sorbi‡.	mapping panels and the CEPH mega artificial chromosome (XAC) clone	in the thymus and appear in the periphery	Late Onset Families
Mike Mullan*, Jeremy Brown*, Fiona Crawford*,	in the APP	R. J. Polinsky**, W. Wasco ⁺⁺ , H. A. R. Da Silva [®] , J. L. Haines ⁺⁺ ,	0.012, 0.00, and 0.031; 106 bp, 0.074, 0.055, 0.022, 0.00, and 0.031; 108 bp, 0.074, 0.055, 0.022, 0.004, 0.0, and	Alzheimer's disease is the most comm	B. Nacmias [‡] , S. Placentini [‡] , L. Amaducci [‡] ,	(results not shown).	despite low expression of surface TCR com- plexes that are devoid of ζ (or n). Under-	
Nick Irving*, Louise James‡, Rebecca Mant	Segregat the physics	M. A. Pericak-Vance [®] , R. E. Tanzi ⁺⁺ , A. D. Roses [®] , P. E. Fraser	0.031; 110 bp, 0.344, 0.343, 0.257, and 0.438; 112	cause of dementia in the elderly. The path	I. Chumakov§, D. Cohen§, L. Lannfelt∥,	The E5-7 cDNA clones detected transcript on northern blots contain	standing the mechanism by which such SP	E. H. Corder, A. M. Saunders, W. J. Strittmatter,
Phillippa Newton*, Karen Rooke*, Penelope Roques*,	21 was an	J. M. Rommens [‡] & P. H. St George-Hyslop [™]	0.055, 0.054, and 0.0; 116 bp, 0.307, 0.265, 0.243,	genic pathway leading to neurodegeneration and AD is not well understood. However,	P. E. Fraser", J. M. Kommens & P. H. St George-Hysion*	several tissues, including most regions	T cells are generated may provide addition- al insight into the mechanism of positive	D. E. Schmechel, P. C. Gaskell, G. W. Small, A. D. Roses,
Chris Talbot*, Margaret Pericak-Vance§, Allen Roses§,	ted by the	* Centre for Research into Neurodegenerative Diseases, Departments of Medicine (Neurology) and Medical Biophysics, L	120 bp, 0.066, 0.067, 0.054, and 0.094; 122 bp,	least some forms of the disease have a gener		ever, in skeletal muscle, cardiac mus gene is expressed at higher levels as	selection in the thymus.	J. L. Haines, M. A. Pericak-Vance*
Robert Williamson", Martin Rossor", Mike Uwen	D21S1, AP	Research Institute, The Hospital for Sick Children, and Department of Molecular and Medical Genetics, University of Te	0.013, 0.004, 0.014, and 0.0; 124 bp, 0.012, 0.012, 0.012, 0.027, and 0.0; 126 bp, 0.008, 0.0, 0.0, and 0.0.	 etiology. For autosomal dominant, early-one (<65 years) AD, causative mutations has 	 Centre for Research into Neurodegenerative Diseases, Departments of Medicine (Neurology) and Medical Biophysics, 	~2.3 and ~2.6 kb (which we have n		The apolipoprotein E type 4 allele (APOE-ϵ4) is genetically associated with the common
* Alzheimer's Disease Research Group, Departments of Biochemistry and	the genetic	M5S 1A8, Canada § Laboratoire de Neurohistologie, Ecole Pratique des Hautes Etudes and U106, INSERM La Salpetriere, 75651 Paris Ceo	 P. Margartte, C. Bonati-Pelle, MC. King, F. Cler- get-Darpoux, Am. J. Hum. Genet. 50, 1231 (1992). 	been identified in the amyloid precursor pr	University of Toronto, Toronto, and Division of Neurology, Department of Medicine. The Toronto Magnitud Toronto.	from that of the 2.7-kb S182 trans	REFERENCES AND NOTES	from 20% to 90% and mean age at onset decreased from 84 to 68 years with increasing
Neurology, St Mary's Hospital Medical School, London W2 1PG, UK	TABLE 1 Th	USL-6 and UO-CNR, 88046 Lamezia Terme, Italy Department of Neurologic and Development of Elevence, visite Montageri 85. Elevence, Italy	 J. Ott, Analysis of Human Genetic Linkage (Johns Hopkins Univ. Press, Baltimore, 1991); L. L. Cavelli- 	tein (APP) gene on chromosome 21 (1), at mutations, that correcte with familial A	Ontario M5S 1A8, Canada	hybridize with S182 probes at high s	 J. J. J. Sussman et al., Cell 52, 85 (1988); R. D. Klausner, J. Lippencott-Schwartz, J. S. Bonifa- 	number of APOE-€4 alleles in 42 families with late onset AD. Thus APOE-€4 gene dose
333 Cedar Street, New Haven, Connecticut 06150, USA	polynic	# Department of Neurology, University of Turin, via Cherasco 15, 10126 Turin, Italy	Storza and MC. King, Am. J. Hum. Genet. 38, 599 (1986).	(FAD) have been identified in S182, a stro	T Research Institute, The Hospital for Sick Children, and Department of Molecular and Medical Genetics,	sequence predicted a polypeptide of	cino, Annu. Rev. Cell Biol. 6, 403 (1990); J. S. Bonifacino, P. Cosson, R. D. Klausner, Cell 63,	is a major risk factor for late onset AD and, in these families, homozygosity for APOE-e4 was virtually sufficient to cause AD by age 80.
§ Duke University Medical Center, Durham, North Carolina NC 27710, USA Departments of Psychological Medicine and Medical Genetics,	Locus	 Clinical Neuropharmacology Section, NINUS, 9000 Rockville Pike, Bethesda, Maryland 20892, USA ** Centre d'Etude Polymorphisme Humaine, 27 Rue Juliette Dodu, 75010, Paris, France 	 Lod scores computed using the phenocopy correc- tion model for D1S479 were Z_{max} = 4.60 (# = 0.10) 	 candidate gene for the chromosome 14 AI loars (2) In addition to these major and 	University of Toronto, Toronto, Ontario M5S 1A8, Canada ¹ Department of Neurology and Psychiatry, University of Florence,	(Fig. 2). BLASTP alignment analyses	 503 (1990). M. Baniyash, P. Garcia-Morales, J. S. Bonifacino. 	
University of Wales College of Medicine, Cardiff CF4 4XN, UK	D21S16	†† Department of Neurology, University of Massachusetts Medical Center, 55 Lake Avenue, Worcester, Massachusetts 0 ‡‡ Molecular Neurogenetics Laboratory and Laboratory of Genetics and Aging. Massachusetts General Hospital. Department 10 Department of Neurogenetics Laboratory and Laboratory of Genetics and Aging. Massachusetts General Hospital. Department 10 Department of Neurogenetics Laboratory and Laboratory of Genetics and Aging. Massachusetts General Hospital.	and Z _{max} = 6.51 (9 = 0.08) with the VG and control allele frequency for the 112-bp allele, respectively.	ices (2). In addition to these major ge	viale Morgagni 85, Florence, Italy S Centre d'Etude Pebrearchites Humains, 27 rue, Iuliatto Dadu	identity = $20-63\%$ over five domains	L. E. Samelson, R. D. Klausner, J. Biol. Chem. 263, 9874 (1988); D. G. Ortoff, C. Ra, S. J. Frank.	
A LOCUS segregating with familial Alzheimer's disease (AD) has	D21513 D2151	Genetics, Harvard Medical School, Boston, Massachusetts 02114, USA Si Broan Alzheimer's Disease Research Center, Duke Liniversity Medical Center, Durbam, North Camlina, 27710, USA	Under this model, the maximum D1S479 lod score for a single family, the B family, was 2.95 (# = 0.09)	E. Levy-Lahad, P. Poorka, J. Oshma, GE. Yu, E. M mens, G. D. Schellenberg, Geriatric Research, Educ	75010, Paris, France	weak homologies to several other tra lar to those previously reported for s	R. D. Klausner, J. P. Kinet, Nature 347, 189 (1990); A. Bauer et al., Proc. Natl. Acad. Sci.	Alzheimer's disease (AD) is a devastating alleles: APOE-62, APOE-63, and APOE-
been mapped to chromosome 21 (ref. 1), close to the amyloid precursor protein (APP) gene ²⁻⁵ . Recombinants between the APP	APP(EcoRI) 021517	Molecular Pathology, Glaxo Research and Development, Greenford Road, Greenford, Middlesex UB6 OHE, UK	using the control allele frequency for the D1S479 112-bo allele. For other markent, marking lod	tion, and Clinical Center (1828), Veterans Affairs Mede Center, 1660 South Columbian Way, Seattle, WA 981	Department of Clinical Neuroscience (Geriatric Medicine), Huddinge Hospital, Karolinska Institute, 14186 Huddinge, Sweden	most striking alignment of E5-1 was	U.S.A. 88, 3842 (1991). 3 A.M. Weissman et al. EMPO / 8, 2651 (1993); P.	neurologic disorder that affects millions of individuals of all races and ethnic back sporadic AD late operat cases have at last
gene and the AD locus have been reported ⁶⁻⁸ which seemed to	D21S156	 Sandoz Research Institute, Sandoz Pharmaceuticals Corporation, 59 Route 10, East Hanover, New Jersey 07936, U MRC Human Genetics Unit, Western General Hospital, Crewe Road, Edinburgh, UK 	scores under the phenocopy correction model were as follower: D19227 Z = 0.92 = 0.20 D15213	USA. W. Wasco, D. M. Romano, W. H. Pettingell, P. D. Jond	To whom correspondence should be addressed	sequence predicted for the S182 pro 63% overall amino-acid sequence ide	A. Irving and A. Weiss, Cell 64, 891 (1991); S. J. Frank et al. Science 249, 174 (1991); S. J.	grounds. Onset before age 60 is infrequent one APOE-e4 compared to 31% of control
exclude it as the site of the mutation causing familial AD. But recent genetic analysis of a large number of AD families has	APP(Bc/1)		$Z_{max} = 0.27, \theta = 0.20, D1S439, Z_{max} = 3.13, \theta = 0.15, D1S225, Z_{max} = 1.40, \theta = 0.29, D1S2450, Z$	S. D. Schmidt, A. C. Crowley, S. Y. Guenette, R. E. Ta Genetics and Aging Unit, Massachusetts General Hos	WE report the cloning of a novel gene (E5-I) encoded on chromo-	show almost complete identity (61-9	Wegener et al., Cell 68, 83 (1992); C. Romeo, M. Amiot, B. Seed, 2bid, p. 889; F. LeTormeo, M.	and caused by either a mutation in the amyloid precursor protein (APP) gene lo- as an important factor in the aviology of
demonstrated that the disease is heterogeneous?. Families with	Allele freq	Some cases of Alzheimer's disease are inherited as an autosomal dominal	= 0.27, $\hat{\theta}$ = 0.30; D15103, Z _{max} = 2.58, $\hat{\theta}$ = 0.06.	⁴ tal, Charlestown, MA 02129, USA. K. Wang, YH. Fu, D. Gales, Darwin Molecular, Both	some 1 which has substantial nucleotide and amino-acid sequence similarity to the S182 gene on chromosome 14a24.3. Mutations,	domains) (Fig. 2). Furthermore, all S182 in subjects with AD3, including	D. Klausner, Proc. Natl. Acad. Sci. U.S.A. 88, 9061 (1001)	cated on chromosome 21 or, more com-
Some families with early-onset AD show linkage to chromosome 21 markers ^{9,10} .	been calculat al. manuscrin	form of Alzheimer's disease to chromosome 14g24.3 We have defined a minimum	cause of the extraordinary computer time necessary with tractitional mathematical Advantage of the extraordinary computer time necessary	WA 98021, USA. E. M. Wijsman, Division of Medical Genetics, Departm	including three new missense mutations in the S182 gene, are asso-	conserved in the putative E5-1 prot	4. P. E. Love, M. L. Tremblay, H. Westphal, Proc.	moniy, by an unidentified gene on chromo- some 14 (1-4). Previous evidence of the that it is correlated with increased ride and
21 markers, but some do not ^{8,9,11} . This has led to the suggestion	morphisms a	region containing the AD3 gene, and isolated at least 19 different transcripts	ulation methods are under development (A. Kong, N. Cryw M. Brigger M. Bala, Capat. Emission of 40, 490	of Medicine and Department of Biostatistics, Universit, Washington, Seattle, WA 98195, USA.	disease (AD) ¹ . Both the E5-1 and the S182 proteins are predicted	share a similar structural organization	 J. G. Orloff, S. J. Frank, F. A. Robey, A. M. Wittename, D. K. Barrank, F. A. Robey, A. M. 	involvement of chromosome 19 in late earlier onset.
that there is non-attenc genetic neterogeneity even within early	from non-det	this region. One of these transcripts (S182) corresponds to a novel gene w	(1993); S. Lin, E. Thompson, E. M. Wijsman, MA J.	T. D. Bird, Division of Neurology, Veterans Affairs Mod Center, Seattle, WA 98108, USA.	to be integral membrane proteins with seven membrane-spanning domains and a large exposed how between the sixth and support	tain at least seven hydrophobic domains and large acidic hydropho	Heissman, H. D. Kausner, J. Biol. Chem. 264, 14812 (1989); YJ. Jin et al., Proc. Natl. Acad. Col. U.S.A. 97 2010 (1999).	finding of an association between AD and early onset families of these two have
2 Present addresses: Department of Respiratory Medicine, St Mary's Hospital Medical School, London W2 1PG LK (A,H); ICI Biotechnology Division, Alderley Park, Macciestield, Churchine, 14((), 13	confidence, [1 informative w	predicted to contain multiple transmembrane domains and resembles an int	Nemi Appz. Nett. Biol. 10, 1 (1955); S. Lin, E. M. Wijsman Am, J. Hum, Genet. 55, A40 (1994); how-	These authors contributed equally to this work.	transmembrane domains. Analysis of the nucleotide sequence of	terminus and between TM6 and TN	Sci. U.S.A. 67, 3319 (1990); L. K. Clayton ef al., ibid. 88, 5202 (1991).	the apolipoprotein E locus (APOE) on chromosome 21 APP mutations and two
§ To whom correspondence should be addressed.	and D21S82.	protein. Five different missense mutations have been found that cosegregate	ever, such meniods cannot yet accommodate age-	compliant unequination should be addressed.	the open reading frame (ORF) of the E5-1 gene led to the discovery of two missense substitutions at conserved amine-acid residues in	similarity between the E5-1 and S182 of alternative splicing of the respect	 Antibody size was generated in a rabbit to a peptide 	chromosome 19 (6-8). APOE has three have the disease state linked to chromo-
704		tamilial Alzheimer's disease. Because these changes occurred in conserved		SCIENCE • VOL 269 • 18 AUGUST 19	affected members of pedigrees with a form of familial AD that has	lysis of cloned E5-1 products of the		SCIENCE • VOL. 261 • 13 AUGUST 1993 921
© 1991 Nature	Publishing Gro	gone, and are not present in normal controls, they are intely to be causative			a later age of onset than the AD3 subtype (50-70 years versus 30-60 years for AD3). These observations imply that the E5-1	after reverse transcription (RT-PCI from several tissues, including sk		
		ALZHEIMER'S disease (AD) is a degenerative disorder of the cen-	ocus put into effect a		gene on chromosome 1 and the S182 gene on chromosome 14q24.3	revealed alternative splicing at nucleo	tides 1,153-1,250, which	
	- L	trai nervous system which causes progressive memory and cogni- tive decline during mid to late adult life, and is accompanied by of known genes on chromosome 14a resi	to AD. Initial studies ulted in their exclusion		functions, and indicates that mutations in conserved residues of	1 protein (Figs 2 and 3). These residue	s share near-identity (31/	
		a wide range of neuropathologic features including extracellular as candidates for the AD3 locus ^{9,11-13} . W	e report the cloning of		E5-1 could also play a role in the genesis of AD. Our results also indicate that still other AD succentibility arms exist	33) with the alternatively spliced 1 TM6→TM7 loop of \$182 (which corr	esidues 257-290 in the espond to exon 9) (Figs	
		Although the actiology of this disease is complex ² , at least three segregating early-onset, autosomal domin	nant AD.		and an out on survey and the second	with the start with the		
	- L	different genetic loci that confer inherited susceptibility to this disease have been identified. The of (112 Con-Ara) allele of Genetic manning of the AD3 loc	cus			deus deus		
	- L	the Apolipoprotein E (ApoE) gene is associated with AD in a After the initial mapping of the AD3 loca	as near the anonymous		a nigo pus	nic ni nic ni nic ni nic nu nic nu nic nu	vascie	
		significant proportion of cases with late onset (>60 years) ^{3,4} . microsatellite markers D14S43 and D1 Mutations in the sene for the B-amyloid precursor protein investigated the segregation of more than	4S53 (refs 9-11), we n 18 additional genetic		EIG 1 Transpointion of the E5-1 homologue - 0 B B First	Phalar Phalar Phalar Phalar Phalar Phalar Phalar Phalar	a a a a a a a a a a a a a a a a a a a	
		(βAPP) have been found in a small number of families (<3% markers from 14q24.3 (Fig. 1a) in a col	lection of 21 pedigrees		was investigated by hybridization of the E5-1	W(K) M M M M M M M M M M M M M M M M M M M	M(K)	
		of cases) with disease onset before 65 years of age ^{**} . A third locus (AD3) has been mapped by genetic linkage studies to chro- 10). Pairwise maximum-likelihood analyse	dominant trait (see ref. es provided substantial		brain regions (a) and several peripheral tis-	25 -	95*	
		mosome 14q24.3, and may account for up to 70% of early-onset cumulative evidence for linkage betwee	n familial Alzheimer's		sues (c). The E5-1 transcript is smaller and 75 -	75	75 -	
	- L	autosomai dominant AD . Although early-onset AD is less common than late-onset AD, the AD3 locus is associated with ever, many of the genetic data supportin	rrs (z≥+11.00). How- g linkage were derived		which is shown hybridized to the same blot		u -	
	- L	the most aggressive form of this disease (onset 30 to 60 years), from six large, early-onset pedigrees (F	AD1 (ref. 14), FAD2		using identical conditions (0) . 24 -	24 -	24 -	
	- L	(ret. 15), rAD3 (rets 16, 17), rAD4 (ret To whom correspondence should be addressed. and 603 (ref. 20)), each of which provi	ided an individual log				And the second sec	
	- L	754 NATURE · VOL	. 375 - 29 JUNE 1995		NATURE · VOL 376 · 31 AUGUST 1995		775	
	- L							

Karran, E., Mercken, M., De Strooper, B., The amyloid cascade hypothesis for Alzheimer's disease: an appraisal for the development of therapeutics. *Nat. Review Drug Discov.* 10, 698–712 (2011)

APP amino acid substitution A673T was found to protect against AD onset and cognitive decline in cognitively healthy elderly individuals



The mutant amino acid is adjacent to the BACE cleavage site, and its presence resulted in an ~40% reduction in the formation of Aβ peptides in vitro



A673V is a pathogenic variant know to increase amyloid deposits



Prime editing enables insertions, deletions, and all possible base-to-base conversions without double-strand breaks



PRNP (prion) prime editing

Protection can be conferred by a G•C to T•A transversion Prime editing: 53% installation of G127V with 1.7% indels

Exciting future directions in interfacing genome editing strategies with in vivo delivery for tackling disease Solving the protein folding problem

There may be many more misfolded proteins that contribute toward disease



UniProt contains **180 million** unique protein sequences;

to date **170,000** structures have been elucidated

Solving the protein folding problem

There may be many more misfolded proteins that contribute toward disease



CASP: the protein solving challenge

PROTEINS: Structure, Function, and Genetics 23:11-1v (1995)

INTRODUCTION

A Large-Scale Experiment to Assess Protein Structure Prediction Methods

Methods for obtaining information about structure from amino acid sequence have apparently been advancing rapidly. But just what can these methods currently deliver? The following papers present the results of a large scale experiment that we have orchestrated to determine the current state of the art in protein structure prediction. We consider that the only way to objectively assess the usefulness of prediction methods is to ensure that predictions are made without any knowledge of the answers. We therefore set out to provide a framework in which a large number of such blind predictions could be made and evaluated. The procedure consisted of three parts: the collection of targets for prediction from the experimental community, the collection of predictions from the modeling community, and the assessment and discussion of the results.

COLLECTING PREDICTION TARGETS

Information was solicited from X-ray crystallographers and NMR spectroscopists about structures that were either expected to be solved shortly or that had been solved already but not discussed in public. Targets were identified through personal contacts, blanket emailing, and appeals at scientific meetings. The collecting and management of prediction targets proved to be a difficult undertaking. In all, information on 33 different proteins was obtained. Some of these were not solved in time for the prediction experiment and some were made public without sufficient notice to the predictors. Finally, one or more predictions were received on 24 of these targets.

CATEGORIES OF PREDICTION

The difficulty of prediction depends on the extent of the relationship of the target protein to already known structures. For this reason, predictions were divided into three types:

1. Comparative modeling: Cases where there is a clear relationship between the sequence of the target protein and one or more known structures. In these circumstances, it is assumed that the tertiary structures are similar, and an initial model may be based on the structure with the most similar sequence. Thus, an approximately correct fold is as-

© 1995 WILEV-LISS INC

sured. The prediction challenge is then in devising techniques that can determine the detailed structural differences between the target and the known related structures. These techniques deal with the alignment of the target sequence on the templates, the best choice of template structure for each part of the chain, small (of the order of 1 or 2 Å) adjustments of main chain position, the orientation of side chains, and the conformation of stretches of chain not related to any of the template structures (the "loops").

2. Threading, or fold identification: Even when there is no detectable sequence relationship between two proteins, they may have closely related folds. Threading techniques attempt to identify the fold a sequence will adopt by considering its fit to each member of a library of known folds. That is, the sequence is "threaded" onto each fold, and the suitability of the interactions thus created is evaluated using a fitness function. This is a relatively new technique, made possible by the rapidly increasing size of the set of known folds. A wide variety of scoring functions and sequence structure alignment methods are currently being developed. The primary challenge is to unambiguously identify an equivalent fold to the target protein in the database, if one exists. In cases where this can be done, subsidiary questions concern how reliable a model based on the fold similarity would be. For example, is the alignment of the target protein sequence on the related structure correct? At this stage, it is probably not possible to produce as detailed a model as in comparative modeling.

3. Ab initio predictions: All methods that do not rely in a direct way on database approaches. The classical view of the structure prediction problem: presented with nothing but a sequence and some knowledge of the interactions between amino acids, predict the three-dimensional fold. Methods include secondary structure prediction, the use of rules about protein topology, lattice based simulations, and molecular dynamics and Monte Carlo methods.

Received May 30, 1995; accepted June 2, 1995. Address reprint requests to John Moult, Center for Advance Research in Biotechnology, University of Maryland Biotechnology Institute, 9600 Gudelsky Drive, Rockville, MD 20850. In 1994, Prof. John Moult and Prof. Krzysztof Fidelis founded **CASP** as a biennial assessment for blind protein structure prediction

Global distance test (GDT)

Ranges from 0 to 100

% proteins within a certain distance from their correct position

>90 GDT is competitive



Google's UK-based lab and research company has solved the problem

Median Free-Modelling Accuracy



CASP



CASP: the protein solving challenge



T1037 / 6vr4 90.7 GDT (RNA polymerase domain)



T1049 / 6y4f 93.3 GDT (adhesin tip) Experimental resultComputational prediction

Requires very modest computing power: ~100-200 GPUs run over a few weeks



Hopefully many more protein structures at risk of misfolding can be identified