

SN: 08/375,887

FD: 1-20-95

PATENTS-US--A8375887

S-75,872

RECEIVED

MAR 17 1997

OSTI

W-31-109-ENG-38

Fred J. Stevens

METHOD FOR PRODUCING FABRICATION
MATERIAL FOR CONSTRUCTING
MICROMETER-SCALED MACHINES,
FABRICATION MATERIAL FOR
MICROMETER-SCALED MACHINES

METHOD FOR PRODUCING FABRICATION MATERIAL FOR
CONSTRUCTING MICROMETER-SCALED MACHINES, FABRICATION
MATERIAL FOR MICROMETER-SCALED MACHINES

Fred J. Stevens

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

MASTER

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

for

375,887

DISCLAIMER

**Portions of this document may be illegible
in electronic image products. Images are
produced from the best available original
document.**

METHOD FOR PRODUCING FABRICATION MATERIAL FOR
CONSTRUCTING MICROMETER-SCALED MACHINES, FABRICATION MATERIAL
FOR MICROMETER-SCALED MACHINES.

CONTRACTUAL ORIGIN OF THE INVENTION

5 The United States Government has rights in this invention pursuant
to Contract No. W-31-109-ENG-38 between the U.S. Department of Energy
and The University of Chicago representing Argonne National Laboratory.

BACKGROUND OF THE INVENTION

1. Field of the Invention

10 This invention relates to nanometer-scaled macromolecules and to a
method for producing micrometer-scaled macromolecules, and more
specifically, this invention relates to a method for producing self-
assembling building blocks for micrometer-scaled machines, and to
nanometer-scaled, self assembled fabrication material for micro-scaled
15 machines.

2. Background of the Invention

The concept of building micrometer-scaled machines is not new. The applications of such microscopic devices are widespread, including microcircuitry, drug delivery, textile production, and building materials.

5 The advantage to using products produced from such building techniques is that said products are built "from the bottom up" to exact specifications, resulting in little or no wastage occurring during the manufacturing process. In addition, the dynamic nature of constructing fabrication materials "atom-by-atom" at equilibrium confers high fidelity and self-
10 repair capabilities.

Biology has served as a starting point for nanotechnology. One such attempt toward building geometrical objects utilizes polynucleotides (U.S. Patent No. 5,278,051). However, and as is true with any new engineering endeavor, design freedom is expanded when choosing from a broader range
15 of parts and chain topologies. Proteins offer this requisite diversity. The 20 amino acids found in standard proteins give proteins greater chemical and structural versatility than is found in structures consisting only of the four nucleotides of DNA or RNA. Further, and unlike DNA and RNA structures, proteins can be dense, compact and relatively stiff.

20 With the greater potential for diversity in protein engineering comes the fold prediction problem. Protein chains do not bind to one another by pair-wise matching of their monomers to arrive at tertiary structures. No simple notion of complementarity can be used to predict or design the folding of a protein. As such, efforts at *de novo* protein engineering
25 remain stymied.

Manipulating already existing protein molecules, essentially as

building blocks or biomodules, can serve as an alternative to *de novo* synthesis. Protein molecules would have to be selected based on their ability to form solids in certain environments.

Most larger proteins contain two or more polypeptide chains between which there are usually no covalent linkages. Instead, the linkages are of a noncovalent nature, and include relatively weaker, less directional bonds such as ionic bonds, hydrogen bonds, and van der Waals interactions. Noncovalent linking processes facilitate self organizing synthesis, a myriad examples of which occur in nature and include molecular crystals, liquid crystals, colloids, micelles, emulsions, phase separated polymers, Langmuir-Blodgett films, self-assembled monolayers, amyloid, viral capsids, among others.

Essential to successful materials fabrication at the microscopic level is the need to control the noncovalent interactions of amino acid side chains with each other, the solvent, and with other solutes. The fabrication process also must be able to overcome the intrinsically unfavorable entropy involved in bringing many molecules together in a single aggregate. Interactions between molecules or parts of molecules must be more favorable energetically than competing interactions with solvent.

In summary, nanotechnology material fabrication techniques must be able to manipulate and build complex structures containing billions of atoms, each atom occupying a predictable location. In turn, the biomodules or "bricks" comprising these structures must be diverse and capable of nearly unlimited manipulation. Heretofore, a fabrication technique to build such complex nanometer scaled devices has remained

elusive.

A need exists in the art for a method to engineer biomolecules so as to aid in the fabrication of building materials to construct first generation molecular machines. The building materials must withstand a myriad of environments to which micro-scaled machines would be subjected. The method also must be dynamic thereby resulting in the fabrication materials assembling and disassembling themselves with error checking ability.

SUMMARY OF THE INVENTION

It is an object of the present invention to provide a method for producing fabrication materials which can be used to construct micro-scale machines that overcomes many of the disadvantages of the prior art.

It is another object of the present invention to provide a method for producing fabrication materials to be used in the construction of micro-scaled machines. A feature of the invention is the utilization of biomolecules which exhibit tendencies for aggregation. An advantage of the invention is the elimination of the fold prediction problem associated with *de novo* protein synthesis.

Yet another object of the present invention is to provide an economical method for fabricating material to be used in the construction of micro-scaled devices and structures. A feature of the invention is constructing fabrication materials from an "atoms up" approach. An advantage of the invention is the elimination of waste that occurs when fabrication materials are produced on a macroscopic scale.

Still another object of the present invention is to provide a dynamic method for producing a fabrication material at equilibrium. A feature of the invention is the predisposition toward self-assembly exhibited by the

fabrication material subunits. An advantage of the invention is a high fidelity characteristic of the assembly process that also results in self repair of the fabrication material.

Another object of the present invention is to provide a fabrication material to be used in the construction of micro-scaled machines. A feature of the invention is the utilization of biomolecules as biomodules having predictable characteristics in buffers of varying pH and ionic values. An advantage of the invention is the incorporation of such fabrication materials in various micro-machines which will be operating in selected environments.

Briefly, the invention provides a method for producing fabrication material for use in the construction of nanometer-scaled machines comprising isolating a first and second protein molecule, each having a surface that contacts other similar molecules to form aggregates in predetermined environments, substituting specific amino acids at predetermined residue positions of the contact surface of the first protein molecule, producing specific quantities of the now substituted first protein molecule, substituting specific amino acids on the contact surface of the second protein molecule at adjacent residue positions to those now-substituted residue positions on the contact surface of the first protein molecule, producing specific quantities of the now substituted second protein molecule, and combining the quantities of the first and second substituted protein molecules so as to facilitate linkage between the molecules, thereby creating a construct of a predetermined size. The invention also discloses a fabrication material for use in the construction of micro-scaled machines comprising a biomodule attached in a predefined

manner to identical biomodules to form a macroscopic structure of predetermined length.

BRIEF DESCRIPTION OF THE DRAWING

These and other objects and advantages of the present invention will become readily apparent upon consideration of the following detailed description and attached drawing, wherein:

FIG. 1 is a series of size-exclusion chromatograms illustrating solution dependence of polymerization properties, in accordance with the features of the present invention;

10 FIG. 2 is a chromatogram showing the effects of solution temperature on aggregating proteins, in accordance with the features of the present invention;

FIG. 3 is an exemplary model for the interaction between light chain dimers that leads to oligomer formation, in accordance with the features of the present invention; and

FIG. 4 is a computer generated image of two Bence Jones dimers with highlighted sterically-favorable residue sites, in accordance with the features of the present invention.

DETAILED DESCRIPTION OF THE INVENTION

20 The invention describes a method for first supplying preformed Bence Jones proteins (native and recombinant) and secondly positioning these proteins to assemble larger macromolecules for use as fabrication materials in the construction of micro-scaled machines.

The protein supply step is initiated by *in vitro* evolution of monoclonal antibody molecules, or subassemblies, manufactured by bacteria. The engineered modules are then assembled either via Brownian

Motion or via atomic force microscopy to formulate larger constructs.

While the proteins or biomodules are the products of sequential covalent processes, the subsequent assembly process relies upon noncovalent self-assembly of the protein modules to obtain the final structure. This differs from some *de novo* efforts which rely on the creation of single, large, covalently-linked structures.

Noncovalent bonds are preferable as a linkage means for a myriad of reasons, including the following:

- 1.) Noncovalent bonds form under conditions that preserve covalent structures;
- 2.) Noncovalent bonds link complementary, polyatomic surfaces so as to enable specific binding; and
- 3.) Noncovalent bonds allow for the equilibration of a system and also allow for the breaking and forming of bonds. This feature provides the advantage of enabling the system to "search" for the specific design configuration sought by the designer.

A large variety of different protein molecules are used in the fabrication process, each protein having a predetermined shape, attachment sites, and selected affinities for other entities. These other entities include viral capsid proteins, synthetic proteins, proteins from the constant domain region of IgG molecules, heavy-chain protein, combinations of these proteins and other proteins, protein aggregates and/or inorganic moieties, such as metals. Topologies of selected proteins can be altered by branched and cyclic structures.

The selected affinities for these entities are controlled so that desired combinations proceed in a predetermined sequence to provide a

predetermined structural shape or layer. In some instances, protein solutions are combined in a predetermined sequence to produce the desired aggregate. The present invention utilizes similar protein types as subassembly modules.

5 Biomolecules that can serve as viable subassemblies or building blocks to more complex fabrication materials include human immunoglobulin light chains. It is well known that certain of these chains, called Bence Jones proteins, form pathological protein deposits when plasma cell dyscrasias result in excess production of the chains and when certain
10 solution conditions are met, including pH values, ionic strength and concentrations. Many different types of proteins are known to form amyloid, but any particular amyloid deposit contains an essentially homogeneous protein core of primarily β -sheet structure. Bence Jones protein is a dimer of light chains, each chain composed of two independent
15 β -barrel domains each in turn consisting of 110-120 amino acids, for a combined molecular weight of approximately 23,000.

Rigid Assembly Minimizes Entropy

Stabilization of the desired structure is crucial. Decreasing the conformational freedom of the backbone of the formulated linear polymer decreases the entropy of unfolding, thus stabilizing the desired folded state. Folding is equivalent to compression; less-flexible molecules
10 expand into a smaller region of configuration space when unfolded, hence the work of compression required to fold them is reduced. For example, in many instances, the compatible replacement of Gly with Ala (or with any amino acid having a β -carbon) or of Ala with Pro each increase stability by $\approx 1\text{kcal/mol}$. Placing an alanine in a crowded location often results in

greater stabilization and rigidity due to noncovalent interaction between the β hydrogens and adjacent structures. In augmented proteins, various cyclic amino and imino acids could be used to constrain conformations in a similar manner, but with varied geometries.

5 There has been great interest in determining the sequences of amyloid-associated immunoglobulin light chains and comparing them to sequences of non-pathogenic light chains to identify regions of the proteins responsible for self-association and fibril formation. A myriad of sequence archives and databases now exist, including those by Kabat
10 E.A. et al. "Sequences of Proteins of Immunological Interest", 5th Ed., National Technical Information Service, U.S. Department of Commerce, GenBank, the European Molecular Biology Lab (EMBL) database, the National Center for Biological Information, and the Brookhaven Protein Database.

 The inventor has searched certain of these archives, particularly
15 those pertaining to the kappa (κ) light chain structures, and found that nearly identical amyloid-forming proteins have similar globular structure with amino acid residue positions that can facilitate noncovalent linkage. The utility of the placement of these sterically favorable amino acids, found either in the native configuration or substituted, is multi-faceted.
20 In some instances, moieties on juxtaposed residue positions have opposite charges, thereby leading to salt bridge formation and subsequent fabrication of larger macromolecules. In another approach, moieties on different molecules but in close proximity to each other make up poly-dentate ligands, with oligomer formation occurring upon addition of metals.

25 Also, many of these proteins bind partially exposed ligands. By using these ligands, or by synthesizing ligand analogs with protruding

moieties, the resulting coordination complexes can be utilized as AFM tips with the protein molecule as the support structure or “handle.”

Exterior groups are more likely than internal moieties to facilitate oligomer formation given that the receptor sites of the interior groups are formed by the rest of the fold. However, suitable R-groups can be substituted to effect folding configurations. As in staged thiol pairing, the time of linkage is controlled by solution manipulation. For example, suitable interior R-groups form constraining cross links when metal ions are introduced into solution. In standard proteins, these constraints force the folding of structures as small as 30 residues.

In standard proteins, regularities at the level of secondary structure provide a partial escape from the protein fold quandary, as do regularities at sterically favorable end regions of elongated protein shapes. Additionally, in augmented proteins, a large R-group could provide a preformed, fold-independent binding site for another R-group, yielding modularity at the level of monomer pairs. The isolation and subsequent positioning of these proteins in relative close proximity so as to induce noncovalent interaction provides the fundamental mechanics for fabrication material building.

While different protein structures have widely different moduli of elasticity, generally, any protein capable of forming Amyloid is a suitable building material. A Young's modulus of ≈ 10 GPa is common to various kinds of polymers of globular protein molecules. (Oosawa, F. and S. Asakura, *Thermodynamics of the Polymerization of Protein*. Academic Press, London, 1975). This falls within the range of moduli of polymeric materials such as polyethylene at 1 GPa, polystyrene at 3 GPa,

polyphenylene sulfide at 8 GPa, poly(chlorotrifluoroethylene) at 14 GPa, and most kinds of wood which range from between 8 and 12 GPa.

Subassembly Construction and Isolation

Biotechnology is currently the most viable technology for the engineering and production of the desired, amyloid-producing light chain
5 dimer construction modules. In the present endeavor, the applicant assessed human immunoglobulin light chains of the K_{IV} subgroup for their aggregation tendencies. All recombinant VK_{IV} domains contain 114 amino acid residues.

Protein engineering is utilized not only to produce natural proteins
10 *in vitro*, but also to incorporate any of the standard amino acids and mono- and/or poly-dentate ligands at sterically favorable positions. An exemplary engineered assembly is depicted in FIG. 3, discussed infra, showing positioned glutamic acid molecules of one dimer in electrostatic linkage with lysine of another dimer to create a larger entity.

15 In addition to facilitating substitutions to invoke the formation of salt bridges, acidic amino acids (aspartic acid, glutamic acid) or basic amino acids (lysine, arginine and histidine) or other mono-dentate ligands can be positioned so as to facilitate the docking of a myriad of metals to facilitate sublinking. These metals include the actinides, such as plutoni-
20 um, neptunium, hafnium, thorium, uranium, other environmentally significant metals, and more common ubiquitous metals including calcium, iron, zinc, magnesium and manganese.

The inventor applied standard recombinant methods to produce myeloma-associated human immunoglobulin light chains in *Escherichia*

coli. The protocol outlined infra provides the basis for observation of the biophysical and biochemical properties of selected human light chain proteins, and mutated versions thereof at particular residues.

Three proteins were isolated: a "benign" protein (LEN) and two amyloid-producing proteins, one (REC) from the urine and a second (SMA) from the lymph nodes of patients suffering from protein deposition diseases. The three-letter designation of the proteins are truncated versions of the names of persons from which the protein samples were drawn.

10 The inventor constructed genes encoding the variable regions of the above three human light chains and have used these synthetic genes to produce the recombinant proteins. The resulting sequences are depicted in Table 1, below. The protein sequence determined for LEN is identical to the previously published sequence (Schneider M., et al "The primary
15 structure of a monoclonal immunoglobulin L-chain of κ -type, subgroup IV (Bence-Jones protein Len): a new subgroup of the κ -type L-chain." *Hoppe-Seyler's Z. Physiol. Chem.* 355:1164-1168) with the exception of position 9 where we find an Asp instead of the Asn previously reported.

Purification of light chain proteins from urine samples was
20 accomplished as described in Solomon A. 1985 "Light chains of human immunoglobulins." *Methods enzymol.* 116:101-121. The proteins were identified as K_{IV} light chains using polyclonal rabbit anti-human K_{IV} antisera, also as described in Solomon, 1985. Amyloid fibril proteins were purified by the Pras method, as described in Pras M., et al. "The
25 characterization of soluble amyloid prepared in water." *J. Clin. Invest.* 47:924-933.

An M13 clone of the germline V-region exon of the human K_{IV} light chain was obtained from ATCC (#61121), and the plasmid pBS^{+/-} was purchased from Stratagene (LaJolla, California). The vector pASK40, which contains an *ompA* signal sequence prior to the polylinker, was provided by A. Skerra, as outlined in Skerra A., et al. "The functional expression of antibody Fv fragments in Escherichia coli: improved vectors and a generally applicable purification technique." *Bio/Technology* 9:273-278. Recombinant proteins utilizing this vector are synthesized as fusions with the *ompA* leader sequence, which is cleaved during transport of the mature protein into the periplasmic space of the host E. coli.

Three strains of E. coli were used as plasmid hosts: DH5 α F- ϕ 80dlacZ Δ M15 (Δ lacZYA-argF) from BRL (Gaithersburg, Maryland) was used for cloning and plasmid preparation; CJ236 [dut1 ung1 thi1 relA] pCJ105(Cmr) from BioRad (Hercules, California) was used to prepare uracil-substituted DNA for site-specific mutagenesis, as described in Kunkel TA., et al. "Rapid and efficient site-specific mutagenesis without phenotypic selection." *Meth. in Enzymol.* 154:367-382. BL26 (F⁻, ompT r_B⁻ m_B⁻ lac) from Novagen Madison, Wisconsin) was used for production of recombinant proteins.

Nucleic Acid Methodologies: Plasmids were prepared and analyzed according to standard methods, per Sambrook J., et al., *Molecular Cloning: A Laboratory Manual*, 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989. Dideoxy sequencing of plasmids was performed with Sequenase kits (United States Biochemical, Cleveland, OH) using standard procedure, such as that outlined in Sanger F., et al. "A rapid method for determining sequences in DNA by primed synthesis with DNA

polymerase. *J. Mol. Biol.* 94:441-448. Muta-Gene *In Vitro* Mutagenesis Kits from BioRad were used for site-specific mutagenesis per the uracil-replacement method outlined by Kunkel, et al., referenced above. Polymerase chain reactions (PCR) as outlined by Saiki RK., et al. "Enzymatic Amplification of β -Globulin Genomic Sequences and Restriction Site Analysis for Diagnosis of Sickle Cell Anemia." *Science* ., 230:1350-1354 were run on a Perkin Elmer thermal cycler 9600 using either Taq DNA polymerase (Perkin Elmer) or Vent DNA polymerase, from New England Biolabs, (Beverly, MA) with temperatures and cycle times determined empirically.

10 Oligonucleotides used for mutagenesis PCR reactions, and plasmid construction were designed using the OLIGO software, as disclosed in Rychlik W., "A computer program for choosing optimal oligonucleotides for filter hybridization, sequencing and *in vitro* amplification of DNA." *Nucl. Acid Res.*, 17:8543-8551, or Genetics Computer Group Sequence Analysis Software Package, as disclosed in Devereux J., "A comprehensive set of sequence analysis programs for the VAX." *Nucl. Acids Res.* 12:387-395.

Construction of plasmids: The BamHI-SphI fragment containing the VKIV exon was subcloned from ATCC clone #61121 into the bifunctional vector pBS+/- and sequenced, confirming the VKIV coding sequence reported in GenBank by Klobeck HG., et al. "Subgroup IV of human immunoglobulin K light chains is encoded by a single germline gene." *Nuc. Acid Res.* 13:6515-6529 and Benson D., et al. GenBank. *Nucl. Acid Res.* 21:2963-2965. The VKIV exon was amplified by PCR with a sense-strand primer which positioned a HincII site at the first codon of the mature KIV light chain (Asp in Table 1) and an antisense-strand primer which added to the

15
20
25

3' end of the VKIV exon the 12 codons of the LEN -segment (Tyr 96-Arg 108 in Fig. 3), tandem stop codons, and a Hind III site. The amplified fragment was digested with HincII and HindIII and cloned into the vector pASK40, which had been digested with EcoRI, blunted with mung bean
5 nuclease, and digested with HindIII. This generated a complete germline VKIV-domain coding region (including both V- and J-segments) following the *ompA* signal sequence of the pASK40 vector, with an additional Ala codon encoded at the junction of the blunted EcoRI site of the vector with the HincII site of the insert. By site-specific mutagenesis, the codon for
10 Asn29 was mutated to Ser to generate the rLEN VKIV-domain. By recombinant PCR, the additional Ala codons at the beginning of the recombinant germline and rLEN VKIV-domain constructs were removed, and a codon for a terminal Arg (position 108 in Fig. 3) was added at the 3' end of the rLEN construct. These final plasmids containing the germline and rLEN VKIV-
15 domain sequences are called pkIVex001 and pkIVprt#1004, respectively.

Since the REC sequence contained 14 differences from LEN, many steps would have been required to derive a gene encoding the REC sequence (SEQ ID NO: 3) using the germline- or rLEN-encoding plasmids as template. Instead, the rREC VKIV coding sequence was synthesized *de novo* using
20 recursive PCR (as outlined in Promodrou C., et al. "Recursive PCR: a novel technique for total gene synthesis." *Prot. Engineering* 5: 827-829) with 8 overlapping oligos of a length of 70-85 nucleotides. The rREC sequence synthesized included the region of the pASK40 vector from the XbaI site preceding the signal sequence to the end of the *ompA* signal sequence plus
25 a coding region for the entire rREC VKIV-domain followed by stop codons

and a HindIII site. Plasmid pkIVrec002 holds the rREC VKIV-domain code.

Table 1: Amino Acid Sequences of κ_{IV} (LEN, SMA, REC) and κ_{II} (CAG) Variable Regions

		1		5		10		15										
5	LEN ¹	Asp	Ile	Val	Met	Thr	Gln	Ser	Pro	Asp	Ser	Leu	Ala	Val	Ser	Leu	Gly	Glu
	SMA ²																	
	REC ³															Pro		
	CAG ⁴	Asn								Leu		Ser		Thr	Pro			
			20			25		a	b	c	d	e	f					
10	LEN	Arg	Ala	Thr	Ile	Asn	Cys	Lys	Ser	Ser	Gln	Ser	Val	Leu	Tyr	Ser	Ser	Asn
	SMA																	
	REC									Asn	Leu	Asp	Ala				Phe	
	CAG	Pro	Ser	Ser	Arg					Leu	Phe	Asp		Ile	Ser			
			30			35				40			45					
15	LEN	Ser	Lys	Asn	Tyr	Leu	Ala	Trp	Tyr	Gln	Gln	Lys	Pro	Gly	Gln	Pro	Pro	Lys
	SMA	Asn	Arg									Leu						
	REC	Asp	Thr		Thr													
	CAG	Gly	Thr		Phe	Asp		Leu							Ser		Gln	
				50			55			60								
20	LEN	Leu	Leu	Ile	Tyr	Trp	Ala	Ser	Thr	Arg	Glu	Ser	Gly	Val	Pro	Asp	Arg	Phe
	SMA																	
	REC							Ser										
	CAG			Leu		Asn	Ala											
			65			70				75								
25	LEN	Ser	Gly	Ser	Gly	Ser	Gly	Thr	Asp	Phe	Thr	Leu	Thr	Ile	Ser	Ser	Leu	Gln
	SMA																	
	REC																	
	CAG										Arg		Arg	Val	Glu			
			80			85				90					95			
30	LEN	Ala	Glu	Asp	Val	Ala	Val	Tyr	Tyr	Cys	Gln	Gln	Tyr	Tyr	Ser	Thr	Pro	Tyr
	SMA										His				His		Gln	
	REC																Pro	
	CAG	Pro			Gly	Ile		Phe		Met		Ala	Arg	Gln	gap	gap	gap	
				100										108				
35	LEN	Ser	Phe	Gly	Gln	Gly	Thr	Lys	Leu	Glu	Ile	Lys	Arg					
	SMA	Thr		*	*	*	*						Leu					
	REC	Thr		Gly				Val										
	CAG	Thr		Gln	Gly	Thr	Lys	Val	Glu	Ile	Lys	Arg						

40 ¹LEN, ²SMA, ³REC, and ⁴CAG correspond to SEQ ID Nos. 1, 2, 3 and 4, respectively, with residues 27a-27f corresponding to residue numbers 28-33 in the SEQ ID listings, infra.

Table 1 depicts the amino acid sequences of the variable regions of the LEN, SMA and REC κ_{IV} proteins and of the CAG κ_{II} protein. Non-amyloidogenic LEN is the reference protein. A blank space indicates that an amino acid residue identical to that of LEN occurs at a particular position. The asterisks at residues 100-103 of the SMA protein indicate positions where the amino acid sequence of SMA could not be unambiguously determined; it is assumed that the SMA sequence (SEQ ID NO: 2) is identical to LEN (SEQ ID NO: 1) at these four positions. Sequence numbering, complementarity-determining regions 1-3 (CDR1, CDR2, and CDR3), and framework regions 1-4 (FR1, FR2, FR3, and FR4) are marked as in Kabat et al. (1991). Lower case letters between residue positions 27 and 28 are used in Table 1 so as to illustrate the similarities and differences between these four sequences and other widely available digitally archived kappa protein sequences. Except for the SEQUENCE ID listings, infra, all residue locations referred to throughout this specification coincide with those depicted in Table 1. CAG contains 110 residues, the other sequences, 114.

The rSMA VKIV-domain construct was generated from the pkIVprt#1004 plasmid using recombinant PCR to convert the three LEN residues Ser29, Lys30, and Pro40 to Asn, Arg, and Leu, respectively (see Table 1). The resulting plasmid was digested with PstI and HindIII to remove codons for the carboxy-terminal 30 residues of the VKIV-domain. A series of four oligos, two sense-strand and two antisense-strand, encoding the carboxy-terminal 30 residues of the rSMA VKIV-domain sequence were annealed and ligated into the PstI- and HindIII-cut plasmid,

giving rise to plasmid pkIVsma007, which contained the entire rSMA VKIV-domain sequence.

The *ompA* signal sequence and VKIV-domain coding sequences of all plasmids were confirmed by dideoxy sequencing.

5 **Purification of recombinant domains:** BL26 cells transformed with plasmids pkIVprt#1004 , pkIVrec002, or pkIVsma007 were grown in 2XYT medium containing 100 µg/ml carbenicillin procured from Sigma Chemical Company (St. Louis, Mo) at 30°C. When the culture reached an A_{595} of 0.75 to 1.0, expression was induced by addition of isopropyl β-D thiogalacto-
10 pyranoside (IPTG; provided by Sigma Chemical Co.) to a final concentration of 1mM. Cell growth was continued for an additional 16 hours. A low agitation rate (100-115 rpm) and 30°C temperature were used throughout the growth period, since growth at faster agitation rates or higher temperatures frequently resulted in cell lysis which contaminated the
15 periplasmic fraction with cytoplasmic proteins.

Preparation of the periplasmic extracts was based on Pluckthun A. et al., "The consequences of stepwise deletions from the signal-processing site of β-lactamase." *J. Biol. Chem.* 262:3951-3957. Cells were harvested by centrifugation at 4000xg for 10 minutes at 4°C.
20 The cell pellet from each liter of culture was gently resuspended in 20 ml ice cold TES buffer (200 mM Tris, 0.5 mM EDTA and 0.5 M sucrose, pH 8.0). One ml of a freshly prepared solution of lysozyme (20 mg/ml in TES) was added to the suspension followed by 40 ml diluted TES (1:1 in H₂O). The cells were incubated on ice for one hour with gentle shaking and then
25 centrifuged at 27,000xg for 15 min. at 4°C. The supernatant constituted

the periplasmic fraction.

The periplasmic fraction from 2 liters starting culture was dialyzed against 10mM Tris, pH 8.0. The volume was reduced by ultrafiltration on an Amicon stirred cell, and the sample was applied at a flow rate of 0.7
5 ml/min to two 5 ml Macrorep Q cartridges, from Bio-Rad, connected in series and equilibrated with the same Tris buffer. At this step, rLEN and rSMA eluted in the flow through fraction with rREC bound to the column and was eluted with a 75 ml 0-100mM NaCl gradient. Fractions containing VKIV domains were identified by SDS-PAGE on Phast Gels (8-25%,
10 Pharmacia), exchanged into 10 mM NaOAc, pH 5.6, by ultrafiltration, and applied at 0.7 ml/min to two 5 ml Macrorep S cartridges (Bio-Rad) connected in series and equilibrated with the same acetate buffer. The rLEN and rSMA proteins eluted from the S-cartridge with a 135 ml 0-150 mM NaCl gradient; the rREC domains eluted in the flow-through fraction.
15 Fractions containing VKIV proteins were exchanged into 20 mM TrisCl, 150 mM NaCl, pH 7.2, concentrated by ultrafiltration, and applied at a flow rate of 0.5ml/min to a HiLoad 16/60 Superdex 75 gel filtration column (Pharmacia) equilibrated with the same buffer. Purified recombinant VKIV proteins appeared to be > 95% pure by SDS-PAGE analysis on 1.5 mm, 13%
20 acrylamide gels using a BioRad Mini-Protean II apparatus with proteins stained with Coomassie brilliant blue. Purified proteins were exchanged into HPLC buffer (20mM potassium phosphate, 100mM NaCl, pH 7.0), concentrated by ultrafiltration to 30-50 mg/ml, and stored at 4°C.

Small-zone size-exclusion HPLC: The HPLC system consisted of a
25 0.3-cm x 25-cm glass column from Alltech Associates (Deerfield, Illinois) packed with Superose 12 from Pharmacia and a Pharmacia 2248 HPLC

pump at a flow rate of 0.06 ml/min. The mobile phase consisted of 20 mM potassium phosphate, 100 mM NaCl, pH 7.0. All experiments were performed at room temperature. Protein samples ranging in concentration from 0.005 to 50.0 mg/ml were injected in a volume of 5 μ l with a
5 Rheodyne 7010 injection valve. The column eluent was monitored at 214, 280, or 254 nm by an HP 1040A diode array detector. Typical run times ranged from 30 to 45 minutes. The data were collected and stored as previously described by the applicant. Stevens FJ., "Analysis of protein-protein interaction by simulation of small-zone size-exclusion chromatog-
10 graphy: application to an antibody-antigen association." *Biochemistry* 25:981-993, and Stevens FJ., "Analysis of protein-protein interaction by simulation of small-zone size exclusion chromatography: stochastic formulation of kinetic rate contributions to observed high-performance liquid chromatography elution characteristics." *Biophys J.* 55: 1155-
15 1167. Chromatograms were normalized by summation of the absorbances at 1000 data points collected during the run and by scaling the data so that the integrated area under the elution profile was equal to 1.

Subassembly Aggregation Detail

Proteins in solution show profound changes in solubility as a func-
20 tion of pH, ionic strength, a solvent's dielectric properties, and temperature. Each unique Bence Jones protein studied for Amyloid formation displayed its own electrolyte characteristics. These characteristics can be exploited via the effects of applied electrical fields, salting out, pH changes, the application of selectively permeable membranes, and other
25 solution manipulations.

As depicted in Table 1, there are few amino acid differences in the

sequences of LEN (SEQ ID NO: 1), REC (SEQ ID NO: 3), and SMA (SEQ ID NO: 2) protein. However, these proteins show marked differences in pathogenic potential. The amyloid-forming proteins REC and SMA differ by 14 and 8 amino acid residues, respectively, from the LEN protein sequence, so differences in pathologic propensities of these κ IV proteins result from the physical-chemical properties associated with the amino acid substitutions.

The inventor has found that the presence of charged residues at position 96 tends to decrease dimerization whereas the presence of hydrophobic residues enhances dimerization. The structure of the dimer is such that positions 96 tend to put two like charges into locations that interact. This is energetically unfavorable and reduces dimerization. By extension, in a mixture of two modified light chains, one of which has a basic residue at position 96 and the other an acid residue, heterologous dimerization would be maximized and homologous dimerization would be minimized.

Another viable residue position for dimerization initiation is position 94 whereby the inventor and colleagues have shown that histidine at that location forms a hydrogen bond with glutamine at position 55 in the second monomer of the dimer. Kolmar H., et al. "Dimerization of Bence Jones proteins: linking the rate of transcription from an Escherichia coli promoter to the association constant of REIv." *Bio. Chem. Hoppe-Seyler* 375:61-70. Positions 55 and 94 therefore are suitably located in a dimer to enable other types of substitutions that would maximize or minimize dimer formation. These residues provide sites for control of homologous and heterologous interactions.

Furthermore, positions 55, 94, and 96 are elements of the complementary determining regions. As such, they directly enable dimerizations with other interactions involved in aggregate (polymer) formation, as discussed infra.

5 Other positions in the framework segments involved in the interface that controls dimerization in naturally occurring light chains include positions 36, 38, 44, 87, 89, and 98. In naturally occurring light chains, the side chains at these positions are involved in symmetrical non-ionic interactions. These positions are highly conserved in natural light chains.
10 Protein engineering at these positions introduce a number of different patterns of ionic interactions which enable heterologous dimerization of certain pairs while inhibiting the interaction of other pairs.

Polymerization Detail

Intermolecular interaction between biomodules that results in
15 polymerization involves several hundred Å² of surface. The total free energy change that accompanies formation of a complex represents the accumulation of numerous favorable and unfavorable contacts, as well as the energetic contribution from removing water from the surfaces involved.

20 **pH Variations:** The inventor has found that the aggregation of purified κ-1 proteins and their modified counterparts proceed at specific pH ranges. A myriad of Bence Jones proteins aggregate at various pHs, as depicted in applicant's co-pending application, Serial No. 8/282,473, the entire contents of which are hereby incorporated by reference.

25 Precipitation pH values are multifaceted in their utility. On the one hand, they can represent solution conditions whereby the interaction of

salt bridges between surface moieties is optimal.

On the other hand, the pH values correspond to those proteins' specific isoelectric pHs, whereby at a specific isoelectric point, the molecule has no net electric charge, and therefore no electrostatic
5 repulsion to prevent aggregation with neighboring and identical protein molecules. Solubility of the protein increases sharply on either side of the isoelectric point, due to the now-ionized species all having the same charge. By first isolating and then solubilizing a particular amyloid protein, the pH of the solution is then adjusted to attain the protein's
10 isoelectric point to allow the solubilized molecules to aggregate. The isoelectric point of certain amyloid forming proteins will dictate the type of pH buffer used for aggregation.

An example of the importance of isoelectric values on the solubility of Bence Jones proteins is the REC protein. Three changes in CDR1 (comple-
15 mentary determining region 1) of that protein result in its isoelectric point being significantly more acidic than that of LEN. With a pI of approximately 3.5, rREC fits the pattern of low isoelectric point being correlated with amyloid potential.

FIG. 1 illustrates the aggregation properties of another Bence Jones
20 protein (CAG). The upper panel demonstrates that the size distribution of the light chain sample is effectively independent of concentration when tested under somewhat neutral (≈ 7.2) pH conditions (50 mM phosphate buffer, 100 mM NaCl). This indicates that the interaction between species of this particular light-chain dimer at this pH is negligible.

25 The lower panel displays the behavior of the same protein under a different set of solution conditions, viz. 30 mM acetate buffer at pH 4.5

(250 mM NaCl). Under these conditions, a shift to a significantly earlier elution position is found, revealing the presence of concentration-dependent aggregation of these light chains induced by the increase in proton concentration in the solvent. The elution position at minute 16 corresponds to a molecular weight appropriate for a light chain tetramer, presumably formed by the association of two dimers; the clear resolution between the apparent tetramer and the lower molecular weight components eluted later suggest that the dissociation rate of the tetramer is slow.

10 Another pH buffer, relatively mid-range at 6.5 viz-a-viz the two disclosed supra, has been employed by the inventor to aggregate proteins with this isoelectric value. The buffer is comprised of 50 mM sodium phosphate, and 0.4 M NaCl.

15 Variable domain fragments of CAG are also capable of forming tetrameric complexes under the same conditions (data not shown).

Electric Field Application: Assembly of protein modules could be facilitated by the application of an electric field to a solution containing protein molecules of primarily the same net surface charge. By allowing the proteins to migrate to the appropriate electrode, and into a predetermined fabrication material pattern (such as a cylinder, sphere, cube, etc.) aggregation can be facilitated by subsequent solution manipulation.

20 Alternatively, salting out with ammonium sulfate or some other neutral salt facilitates precipitation essentially by removing the water of hydration from the protein. Ammonium sulfate is a particularly good candidate given its high solubility in water, thereby resulting in very high ionic strengths.

Temperature Variation: The inventor also has found that certain Bence Jones proteins have characteristic aggregation tendencies at various temperatures. As depicted in FIG. 2, a slightly cast-forming light chain (DOOL) displayed significant aggregation when the solution temperature was raised from 24°C to 50°C. The protein displayed a lower elution volume to total volume (V_e/V_t) ratio at higher temperatures than at relatively lower temperatures.

The chromatogram for protein DOOL shows widely varying solubility characteristics for the protein at between room temperature and body temperature, and between body temperature and 50°C.

Depending on the polymerization initiation step used, what ultimately results is the fabrication of materials ranging in length from approximately 5 nanometers to 1000 microns, and more preferably between 50 microns and 200 microns. For example, the inventor has found that the CAG protein having the chromatogram depicted in FIG. 1 forms large complexes when present in solution at very high concentrations. At the highest protein concentration of this protein, aggregates elute at approximately minute 10. Given the molecular weight exclusion limit of the Superose 12 column of 200,000-300,000, this indicates that polymers comprising at least five light dimers are present in a non-dissociating complex during the ten minute column transit time.

Example 1

FIG. 3 depicts interactions between two identical dimers (A and B) of light chain variable domains. Each dimer consists of two very similar domains consisting of approximately 110 amino acids. Dimers A and B are aligned in the same direction but are related by an approximate 90-degree

rotation about their longitudinal axes. In this illustration, positively charged lysine residues at position 42 are able to interact with negatively charged glutamic acid residues at positions 50, thereby forming two rigid salt bridges.

5 In other examples of amyloid-forming light chains, position 31 rather than position 50, is occupied by an acidic residue, enabling similar double salt bridge formation. When both positions 31 and 50 are occupied by acid residues, a stronger ionic interaction occurs between the dimers. The formation of two salt bridges between adjacent dimers can contribute
10 approximately 4 kcal/mol of free energy change, thereby improving the affinity of interaction by approximately three orders of magnitude compared to light chains in which oppositely charged residues are not found at these locations.

Hydrogen bonding also results in dimer-dimer interaction of the type
15 depicted in FIG. 3. For example, in some cases, positions 42, 31 and/or 50 are occupied by neutral residues, such as asparagine and glutamine. These combinations enable bonding between adjacent dimers.

Hydrophobic substitutions also contribute to the free energy of dimer-dimer interaction. In FIG. 3 for example, the insertion of hydropho-
20 bic residues at positions 30, 40, 49, 91, 94 and 96 correlates with amyloid formation by human kappa light chains. The inventor has found that where a hydrophobic residue is located at position 40 in any human light chain (kappa or lambda), the associated protein self-assembled *in vivo* to form amyloid. In the model shown in FIG. 3, these hydrophobic
25 positions would be buried in the formation of dimer-dimer complexes, an energetically favorable transition.

Numerous other complementary contacts can be engineered into the multiple sites of contact. Amino acid substitutions also can be introduced such that dimers A and B each contribute portions of a tri-dentate or tetra-dentate metal binding site; in this manner, polymerization could be
5 controlled by the addition of a specific metal.

As the affinity constants for the specific metals mentioned above vary, the needed concentrations of the metals in solution also will vary. For example, whereas the affinity for iron at the transfer binding site of hemoglobin is very high at approximately 10^{-24} mols, other metals require
10 micro-molar concentrations to effect linkages. An example of low metal affinity values is the 1mM concentration of the divalent cation Mg^{2+} or Mn^{2+} necessary to restore enolase activity, as reported by the applicant and colleagues. M.J. Peak, et al. "The Hyperthermophilic Glycolytic Enzyme Enolase in the Archaeon, *Pyrococcus furiosus*: Comparison with Mesophilic Enolases." *Arch. of Biochem. and Biophys.* 313, 2, 280-286.
15

Regarding the exemplary dimer interaction depicted in FIG. 3, additional dimers can be added, or stacked, by the same rotational process at either end of the complex.

Other examples of stackable biomolecules include the REC and SMA
20 structures, the three dimensional structures of which are depicted in FIG. 4. First, purified LEN V κ IV protein was crystallized and its structure was determined by x-ray diffraction using the molecular replacement method with light chain V κ I REI as the search molecule. Epp O., et al., "The molecular structure of a dimer composed of the Bence-Jones protein REI
25 refined at 2.0-Å resolution." *Biochemistry* 14:4943-4952.

Using backbone coordinates of the LEN V κ IV dimer, the inventor

generated 3-D structures of REC (above) and SMA (below) by substituting REC and SMA amino acids at positions which differ from the LEN backbone. FIG. 4 illustrates that the majority of differences from LEN are localized on the end of the light chain dimer where the dimer's six CDRs form a pocket, while a single amino acid substitution occurs in each protein at the opposite end of the dimer.

In each of the pathogenic proteins, REC and SMA, an exposed hydrophobic amino acid probably contributes to the protein's amyloidogenic character by decreasing its solubility and/or increasing its tendency to aggregate. Position 28 of REC (in CDR1) contains a Phe instead of the polar residue Asn seen in LEN, while Leu at position 40 of SMA (in FR2) takes the place of Pro in LEN; these positions are indicated with small arrows in FIG. 4. A thick vertical arrow in FIG. 4 points to Pro 96, the only position at the dimer interface of REC which differs from LEN, indicating that this residue plays a role in rREC's increased self-association relative to LEN. Virtually all of the substitutions occur at the end of the dimer where the six CDR loops form a pocket. In REC, the substitution which occurs on the opposite end of the dimer indicates the position at which Pro 15 replaces Leu in LEN.

While the invention has been described with reference to details of the illustrated embodiment, these details are not intended to limit the scope of the invention as defined in the appended claims. For example, aside from using proteins constructed with the 20 basic amino acids, the protein structures disclosed herein can be augmented with nonstandard amino acids, such as those used to facilitate phosphorylation, glycosylation, and other enzyme reactions. Furthermore, extensions of

monoclonal antibody technology can rapidly generate proteins able to bind to almost any selected small molecule or support substrate. All references mentioned in this application are incorporated herein by reference.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Stevens, Fred J

5 (ii) TITLE OF INVENTION: Method for Producing Fabrication Material for Constructing Micrometer-Scaled Machines; Fabrication Material for Micrometer-Scaled Machines.

(iii) NUMBER OF SEQUENCES: Four

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE:

10 Office of Chief Counsel--U.S. Department of Energy

(B) STREET: 9800 South Cass Avenue

(C) CITY: Argonne

(D) STATE: Illinois

(E) COUNTRY: United States

15 (F) ZIP: 60439

(v) COMPUTER READABLE FORM:

(A): MEDIUM TYPE: 3.50 inch, 1.4 MB storage

(B): COMPUTER: Macintosh

(C): OPERATING SYSTEM: Macintosh 7.1

20 (D): SOFTWARE: Wordperfect

(vi) CURRENT APPLICATION DATA: Not Available

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: US 08/282,473

(B) FILING DATE: 29-JUL-1994

25 (viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Harney, Timothy L.

(B) REGISTRATION NUMBER: P-38,174

(C) REFERENCE/DOCKET NUMBER: S-75,872

(ix) TELECOMMUNICATION INFORMATION:

30 (A) TELEPHONE: (708) 252-2042

(B) TELEFAX: (708)252-2183

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 114 amino acids

(B) TYPE: amino acid

5 (C) STRANDEDNESS: Not Applicable

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(ix) FEATURE:

10 (A) NAME/KEY: "LEN" κ_{IV} human immunoglobulin light chain

(B) LOCATION: 1-114

(C) IDENTIFICATION METHOD: Established consensus sequence

(D) OTHER INFORMATION: Expresses physiologically nonaggregating protein.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

15 Asp Ile Val Met Thr Gln Ser Pro Asp Ser Leu Ala Val Ser Leu Gly
 1 5 10 15

Glu Arg Ala Thr Ile Asn Cys Lys Ser Ser Gln Ser Val Leu Tyr Ser
 20 25 30

20 Ser Asn Ser Lys Asn Tyr Leu Ala Trp Tyr Gln Gln Lys Pro Gly Gln
 35 40 45

Pro Pro Lys Leu Leu Ile Tyr Trp Ala Ser Thr Arg Glu Ser Gly Val
 50 55 60

Pro Asp Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr
 65 70 75 80

25 Ile Ser Ser Leu Gln Ala Glu Asp Val Ala Val Tyr Tyr Cys Gln Gln
 85 90 95

Tyr Tyr Ser Thr Pro Tyr Ser Phe Gly Gln Gly Thr Lys Leu Glu Ile
 100 105 110

Lys Arg

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 114 amino acids

(B) TYPE: amino acid

5 (C) STRANDEDNESS: Not Applicable

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(ix) FEATURE:

10 (A) NAME/KEY: "SMA" κ_{IV} human immunoglobulin light chain

(B) LOCATION: 1-114

(C) IDENTIFICATION METHOD: Similarity to related protein consensus sequence.

(D) OTHER INFORMATION: Expresses amyloid-forming protein.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

15 Asp Ile Val Met Thr Gln Ser Pro Asp Ser Leu Ala Val Ser Leu Gly
 1 5 10 15

Glu Arg Ala Thr Ile Asn Cys Lys Ser Ser Gln Ser Val Leu Tyr Ser
 20 25 30

20 Ser Asn Asn Arg Asn Tyr Leu Ala Trp Tyr Gln Gln Lys Leu Gly Gln
 35 40 45

Pro Pro Lys Leu Leu Ile Tyr Trp Ala Ser Thr Arg Glu Ser Gly Val
 50 55 60

Pro Asp Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr
 65 70 75 80

25 Ile Ser Ser Leu Gln Ala Glu Asp Val Ala Val Tyr Tyr Cys Gln His
 85 90 95

Tyr Tyr Ser His Pro Gln Thr Phe Gly Gln Gly Thr Lys Leu Glu Leu
 100 105 110

Lys Arg

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 114 amino acids

(B) TYPE: amino acid

5 (C) STRANDEDNESS: Not Applicable

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(ix) FEATURE:

10 (A) NAME/KEY: "REC" κ_{IV} human immunoglobulin light chain

(B) LOCATION: 1-114

(C) IDENTIFICATION METHOD: Similarity to related protein consensus sequence

(D) OTHER INFORMATION: Expresses amyloid-forming protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

15 Asp Ile Val Met Thr Gln Ser Pro Asp Ser Leu Ala Val Ser Pro Gly
 1 5 10 15

Glu Arg Ala Thr Ile Asn Cys Lys Ser Ser Gln Asn Leu Leu Asp Ala
 20 25 30

20 Ser Phe Asp Thr Asn Thr Leu Ala Trp Tyr Gln Gln Lys Pro Gly Gln
 35 40 45

Pro Pro Lys Leu Leu Ile Tyr Trp Ala Ser Ser Arg Glu Ser Gly Val
 50 55 60

Pro Asp Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr
 65 70 75 80

25 Ile Ser Ser Leu Gln Ala Glu Asp Val Ala Val Tyr Tyr Cys Gln Gln
 85 90 95

Tyr Tyr Ser Thr Pro Pro Thr Phe Gly Gly Gly Thr Lys Val Glu Ile
 100 105 110

Lys Arg

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 110 amino acids

(B) TYPE: amino acid

5 (C) STRANDEDNESS: Not Applicable

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(ix) FEATURE:

10 (A) NAME/KEY: "CAG" κ_{II} human immunoglobulin light chain

(B) LOCATION: 1-110

(C) IDENTIFICATION METHOD: Similarity to related protein consensus sequence.

(D) OTHER INFORMATION: Expresses cast-forming protein.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

15 Asn Ile Val Met Thr Gln Ser Pro Leu Ser Leu Ser Val Thr Pro Gly
 1 5 10 15

Glu Pro Ala Ser Ile Ser Cys Arg Ser Ser Gln Ser Leu Phe Asp Ser
 20 25 30

20 Ile Ser Gly Thr Asn Phe Leu Asp Trp Tyr Leu Gln Lys Pro Gly Gln
 35 40 45

Ser Pro Gln Leu Leu Ile Tyr Leu Ala Ser Asn Arg Ala Ser Gly Val
 50 55 60

Pro Asp Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe Thr Leu Arg
 65 70 75 80

25 Ile Ser Arg Val Glu Pro Glu Asp Val Gly Ile Tyr Phe Cys Met Gln
 85 90 95

Ala Arg Gln Xaa Xaa Xaa Thr Phe Gln Gly Thr Lys Val Glu Ile Lys
 100 105

30 Arg
 110

ABSTRACT

A method for producing fabrication material for use in the construction of nanometer-scaled machines is provided whereby similar protein molecules are isolated and manipulated at predetermined residue positions so as to facilitate noncovalent interaction, but without compromising the folding configuration or native structure of the original protein biomodules. A fabrication material is also provided consisting of biomodules systematically constructed and arranged at specific solution parameters.

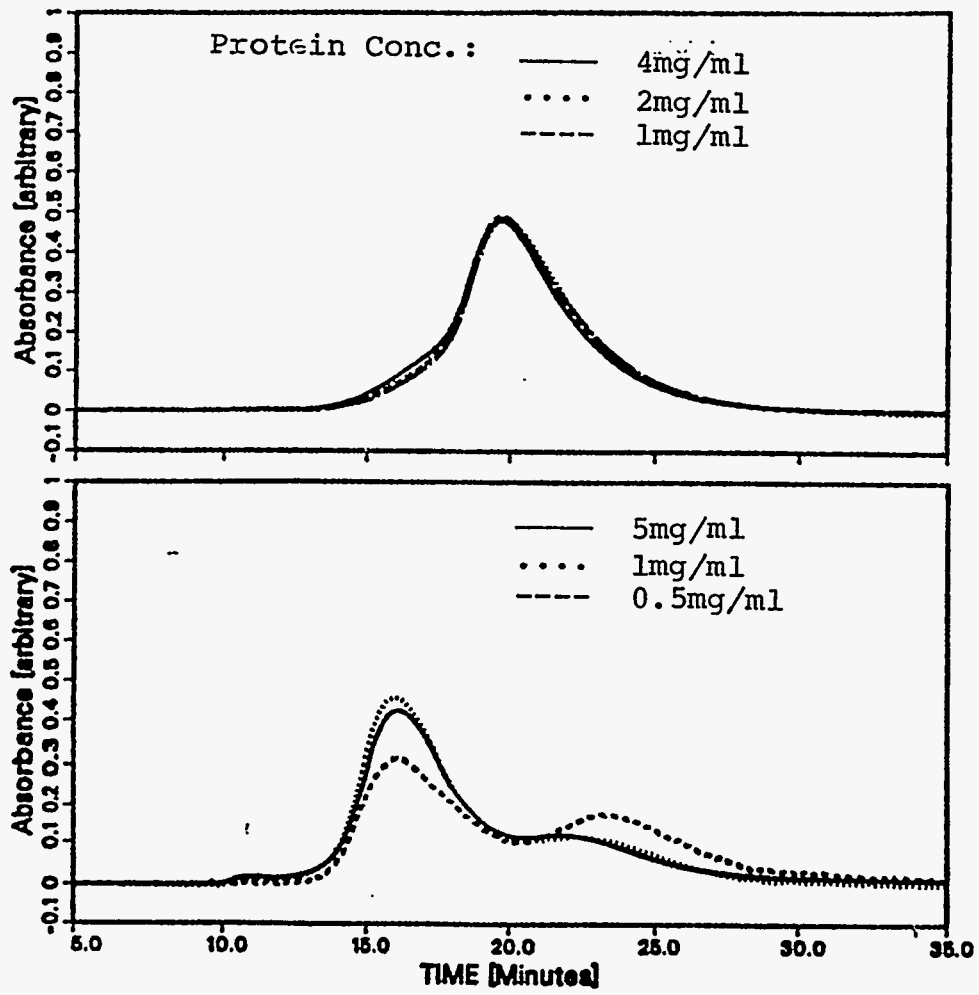


FIGURE 1

Protein Pool

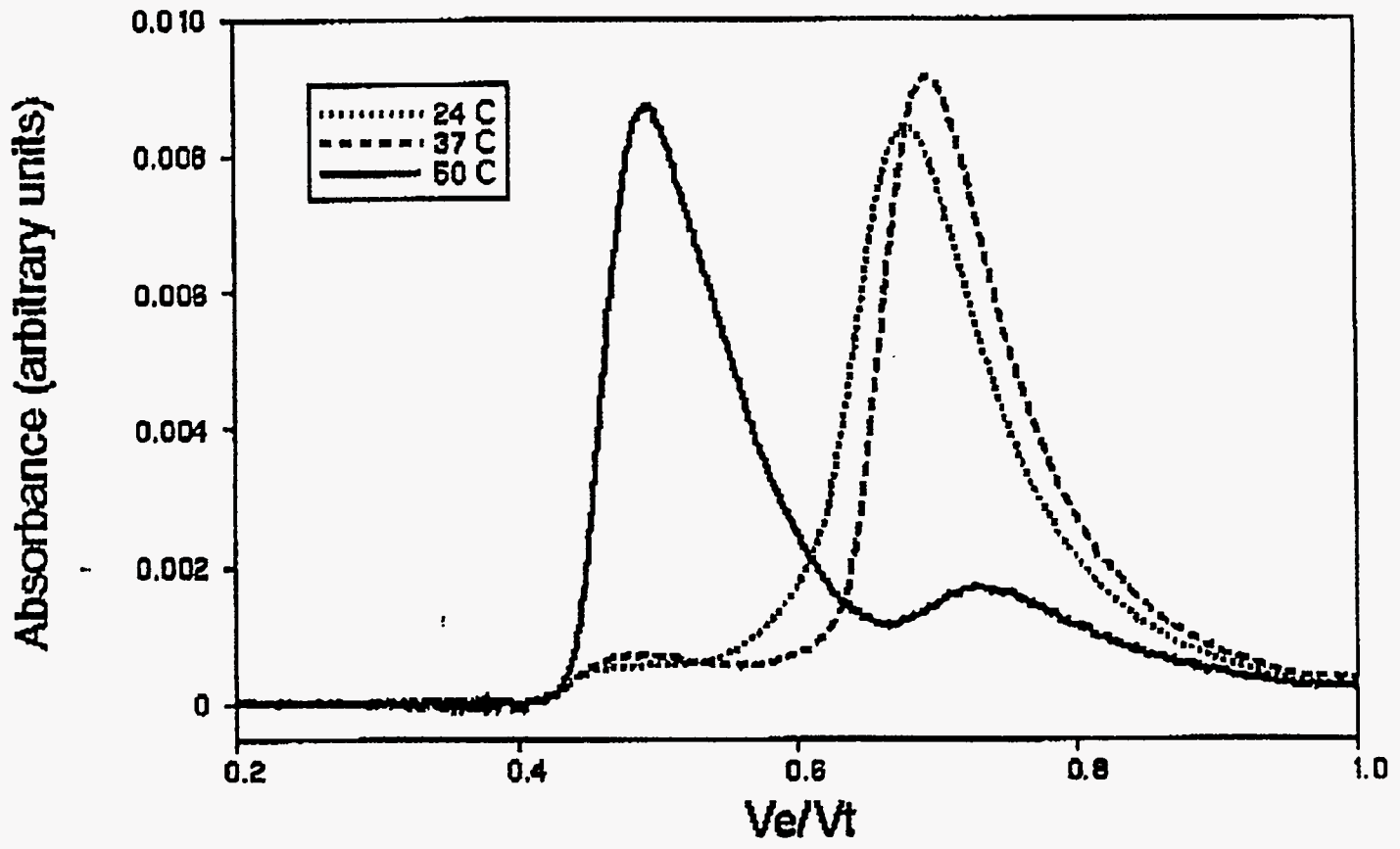


FIG. 2

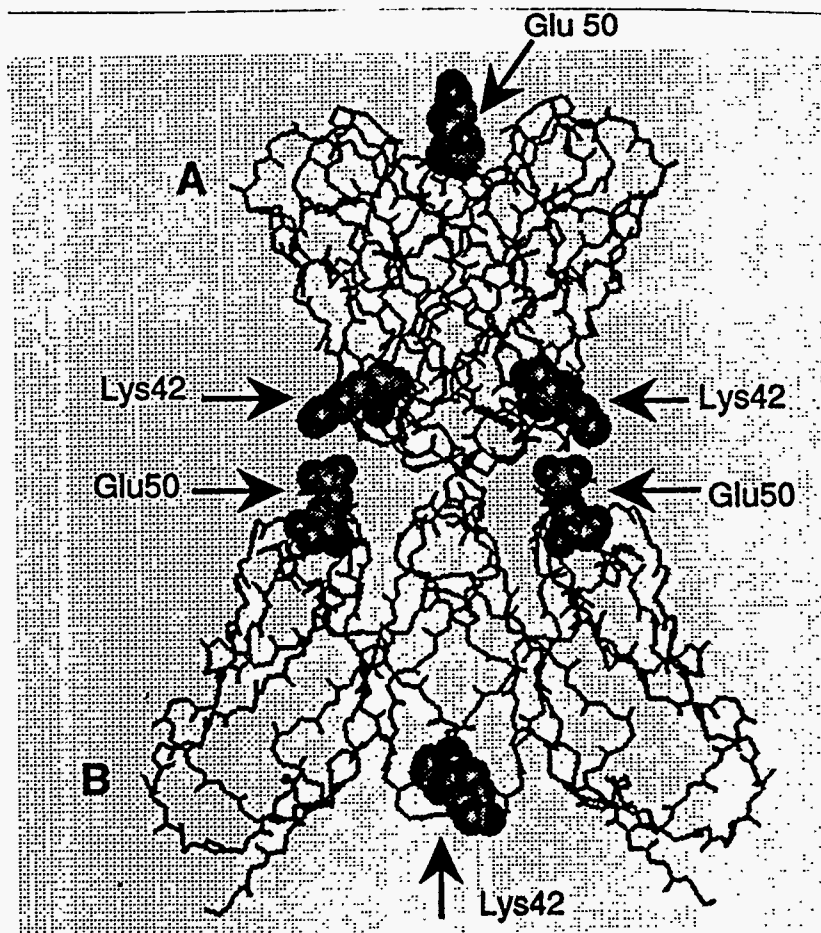


FIG. 3

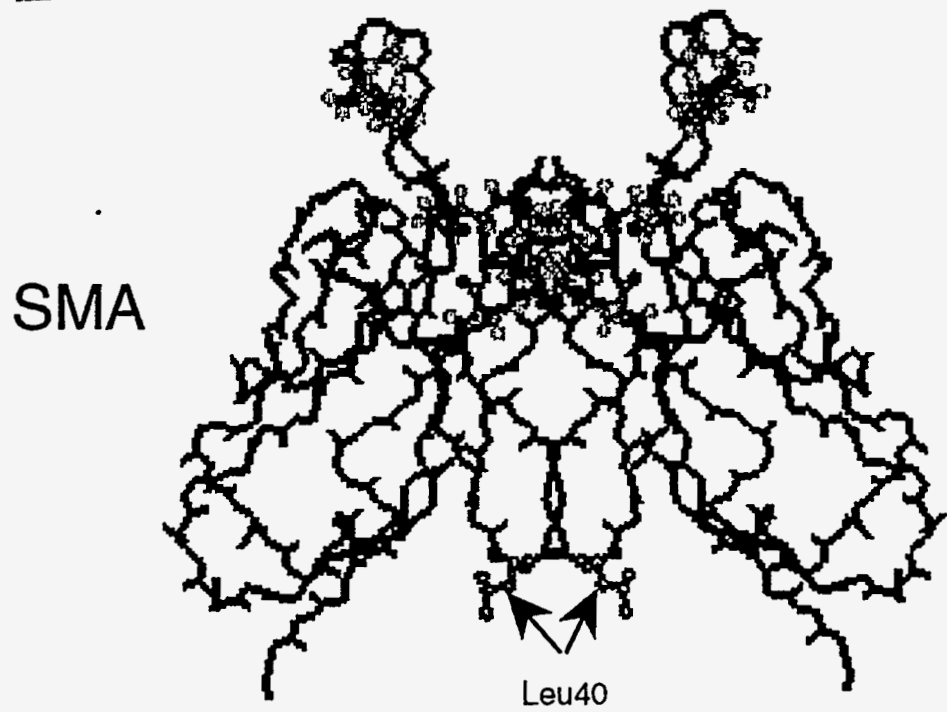
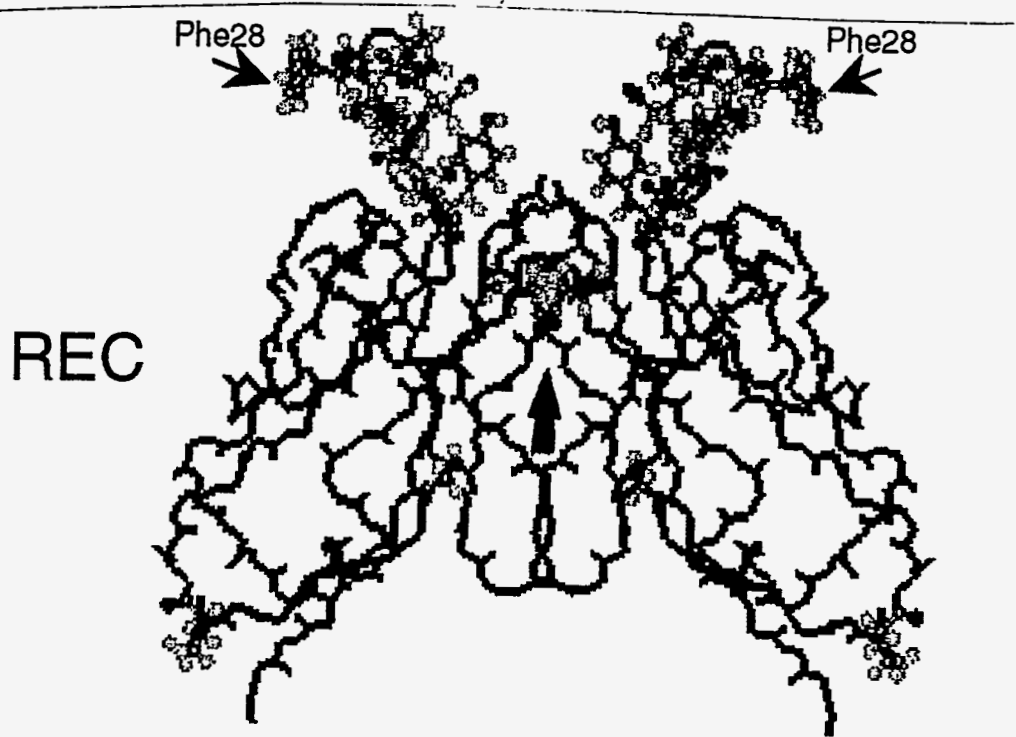


FIG. 4