Volume 38, Issue 6 2010

CRITICAL REVIEWS™ IN BIOMEDICAL ENGINEERING

EDITOR-IN-CHIEF
MARKAD V. KAMATH

begell house, inc.
publishers

Critical Reviews™ in Biomedical Engineering (ISSN 0278-940X) is published bimonthly and owned by Begell House, Inc., 50 Cross Highway, Redding, Connecticut, 06896.

Subscriptions: United States rate for 2010 is $1,759.00. For orders outside the United State and Canada please add an additional $10.00 per issue for foreign airmail shipping and handling fees. Personal (individual) subscriptions must be paid for by personal check or credit card. All subscriptions are payable in advance. Subscriptions are entered on an annual basis, i.e., January to December. For immediate service and charge card sales, call Begell House at (203) 938-1300 Monday through Friday, 9 A.M.-5 P.M. EST. To order by fax: (203) 938-1304. Send written orders to Begell House, Inc., Subscriptions Department, 50 Cross Highway, Redding, Connecticut, 06896. This journal contains information from authentic and highly regarded sources. Reprinted material is quoted with permission, and sources are indicated. A wide variety of references are listed. Reasonable efforts have been made to publish reliable data and information, but the editor and the publisher assume no responsibility for any statements of fact or opinion expressed in the published papers or in the advertisements.

Critical Reviews™ in Biomedical Engineering is covered in *Current Contents~/Life Sciences*, the Scisearch® online database, the *Research Alert*™ Service, and the *Biomedical Engineering Citation Index*™*;* abstracted and indexed in the BIOSYS Database, Elsevier BIOBASE/ *Current Awareness in Biological Sciences*, EMBASE/*Excerpta Medica*, The *Engineering Index* and *COMPENDEX*.

Printed February 17, 2011

CRITICAL REVIEWS™ IN

BIOMEDICAL ENGINEERING

EDITOR

Markad V. Kamath, PhD, P. Eng.
Associate Professor, Department of Medicine
McMaster University, Hamilton, L8N 3Z5, ON Canada
Email: kamathm@univmail.cis.mcmaster.ca

EDITORIAL ADVISORY BOARD

AIMS & SCOPE

Biomedical engineering has been characterized as the application of concepts drawn from engineering, computing, communications, mathematics, and the physical sciences to scientific and applied problems in the field of medicine and biology. Concepts and methodologies in biomedical engineering extend throughout the medical and biological sciences. This journal attempts to critically review a wide range of research and applied activities in the field. More often than not, topics chosen for inclusion are concerned with research and practice issues of current interest. Experts writing each review bring together current knowledge and historical information that has led to the current state-of-the-art.

Each issue contains one or more critical reviews of specified topics representing applied, clinical and basic science areas. Most articles contain in-depth appraisals of the current state-of-the-art in a specific area of research or practice and provide complete and up-to-date bibliographies. Each review attempts to be nearly exhaustive in a constrained area rather than broad and overarching. The critical evaluations of current research and development issues include interpretive discussions of major problems. From time to time, a series of articles in a related topic area are published in order to give comprehensive coverage. Collaborative works generated by multiple authors are frequently used to provide in-depth coverage from multiple viewpoints. Each article is reviewed by one or more independent experts in the field.

The editor invites comments and suggestions about the contents of the reviews.

Markad V. Kamath
Editor-in-Chief

Critical Reviews™ in

Biomedical Engineering

Volume 38 / Issue 6 2010

TABLE OF CONTENTS

Critical Reviews™ in Biomedical Engineering, 38(6):487–509 (2010)

Responding to Change: Thermo- and Photo-responsive Polymers as Unique Biomaterials

Laura A. Wells,[1] Frances Lasowski,[2] Scott D. Fitzpatrick,[2] & Heather Sheardown[1,2]*

[1]Department of Chemical Engineering and [2]School of Biomedical Engineering, McMaster University, Hamilton, Ontario, Canada

*Address all correspondence to Heather Sheardown, Department of Chemical Engineering, John Hodgins Engineering Building,. 1280 Main St. W, Room. 374, Hamilton, Ontario L8S 4L7; Tel.: 905-525-9140 ext: 24794; Fax: 905-521-1350; sheardow@mcmaster.ca.

ABSTRACT: Responsive polymer systems that react to thermal and light stimuli have been a focus in the biomaterials literature because they have the potential to be less invasive than currently available materials and may perform well in the in vivo environment. Natural and synthetic polymer systems created to exhibit a temperature-sensitive phase transition lead to in situ forming hydrogels that can be degradable or non-degradable. These systems typically yield physical gels whose properties can be manipulated to accommodate specific applications while requiring no additional solvents or cross-linkers. Photo-responsive isomerization, dimerization, degradation, and triggered processes that are reversible and irreversible may be used to create unique gel, micelle, liposome, and surface-modified polymer systems. Unique wavelengths induce photo-chemical reactions of polymer-bound chromophores to alter the bulk properties of polymer systems. The properties of both thermo- and photo-responsive polymer systems may be taken advantage of to control drug delivery, protein binding, and tissue scaffold architectures. Systems that respond to both thermo- and photo-stimuli will also be discussed because their multi-responsive properties hold the potential to create unique biomaterials.

KEY WORDS: stimuli responsive; intelligent materials; thermo-responsive; light responsive; drug delivery

ABBREVIATIONS

polyNIPAAm, poly(N-isopropylacrylamide); **LCST**, lower critical solution temperature; **PEG**, poly(ethylene glycol); **GP**, glycerol phosphate; **FGF-2**, fibroblast growth factor; **PLGA**, poly(D,L-lactic acid-co-glycolic acid); **PCL**, polycaprolactone; **HPMC**, hydroxypropyl methylcellulose

I. INTRODUCTION

With the emergence of novel drug therapies, there has been significant emphasis in recent years on the development of adequate delivery strategies to harness the full therapeutic potential of these new pharmaceuticals. Traditional drug-delivery approaches result in an initial spike in drug concentration followed by a steady decline to ineffective levels requiring repeat administration.[1] Commercially available controlled-release devices offer several advantages: they maintain drug concentrations within the desired therapeutic window for prolonged periods of time, stabilize drugs from being degraded by the body, and improve patient compliance.[1] However, many of these polymer-based controlled-release biomaterials are designed to passively deliver drugs in a preprogrammed, unalterable fashion and are unable to adapt to environmental cues. Therefore, stimuli-responsive polymers, also referred to as "intelligent," "smart," or "environmentally responsive" materials, have gained considerable interest for biomedical applications in recent years, particularly in the field of drug delivery. This is attributed to their potential to less invasively provide localized, sustained release of pharmaceutical agents that can respond to internal stimuli in their in vivo environment or to applied external stimuli, as summarized in Table 1.

Stimuli-responsive materials undergo rapid, reversible property transformations in direct response to an environmental stimulus. Such property alterations can include changes in conformation, solubility, and optical properties, as well as alterations of the hydrophilic or hydrophobic balance and surface responses.[2] These environmentally mediated property changes can be utilized to control the biological

1040-8401/10/$35.00

TABLE 1. Internal and External Stimuli Responses That May Be Introduced into Polymer Biomaterials Systems

Stimulus	Mechanism
Temperature	Internal/external
Light	External
pH	Internal
Ionic	Internal
Enzyme	Internal
Magnetic	External
Electrical	External
Ultrasound	External

response of the materials in order to influence drug release rates, degradation kinetics, surface interactions, material conformation, and a variety of other responses.[2]

Internal stimuli are changes that occur naturally in vivo. For example, the pH of various tissues and intracellular compartments varies from 5.0 in intracellular lysozomes to 7.45 in blood.[2] In some cancers, tumor pH can decrease to 6.5[1] or certain enzymes may be present that may cleave polymer bonds that can be used as stimuli for targeted delivery.[3] If materials are injected into tissue, the temperature change to 37ºC may also be a stimulus to cause in situ gelation.[1]

Biomaterials that respond to internal stimuli can also potentially be designed to respond to disease and to natural biological rhythms.[4] External stimuli are those that can be applied from an exogenous source to control biomaterial properties and drug delivery. Light, for example, may be particularly suited to the eye, where it may penetrate to reach the tissue of interest.[5] Magnetic, electrical, and ultrasound stimuli may be applied to induce changes in target tissues to minimize systemic effects.[6] Temperature may also be externally introduced and has been investigated by cycling patients through a hypothermic state.[3] External stimuli are frequently referred to as pulsatile systems because the stimuli are present in a discontinuous state.

Because temperature and light are both minimally invasive stimuli that have been extensively examined for the creation of unique polymer materials for use in a number of biomedical applications, the remainder of this review will focus on thermo- and photo-responsive biomaterials. Photo-responsive polymers in particular have been reviewed less frequently, and the combination of thermo- and photosensitive polymers presents unique opportunities for development. First, thermo-responsive polymers will be briefly explored, followed by a more extensive review of photo-responsive polymers, and finally a review of thermo- and photo-responsive combination polymers.

II. THERMO-SENSITIVE MATERIALS

Recently, injectable, in situ forming polymers have been investigated as a method with which to improve drug and cell delivery.[7] These materials are attractive because the injectable fluid can be introduced into the body in a minimally invasive manner prior to gelling, and therefore surgical procedures are not required for placement.[8] In situ forming systems may also be useful for the treatment of conditions that are not easily accessible. These materials have the potential to reduce surgical side effects and increase patient compliance,[9] thereby improving clinical outcomes. The in situ gelling systems examined in this report consist of polymer solutions that can be prepared at room temperature and subsequently transition from a soluble liquid to an insoluble gel following injection in response to the increased temperature. We discuss current thermo-sensitive materials being explored and their potential clinical uses in the fields of drug delivery and tissue engineering. These examples will illustrate the potential of temperature-sensitive, injectable materials as improved drug- and cell-delivery systems.

The ideal injectable material for biomedical applications is a stable, low-viscosity solution prior to administration to ensure injectability through a small needle for patient comfort.[10] It should have a subphysiological gel-transition temperature and its gelling kinetics need to be rapid enough to ensure entrapment of the therapeutic agent at the target or injection site.[7] Slow-gelling kinetics would allow a large initial burst release of the payload, resulting in potential off-target complications and a decreased concentration of the therapeutic agent at the site of interest. Therefore, gelation of the intelligent scaf-

fold should occur almost instantly upon contacting its stimulus. The material must be compatible with the injection environment[10] and it should also demonstrate a sustained, controlled drug-release profile with the ability to add therapeutic agents through simple mixing.[9] For cell-delivery applications, the material should demonstrate adequate cellular compatibility, particularly the ability to maintain cell adhesion, mimicking the natural extracellular matrix.[7] Fulfillment of these requirements allows the material to be used effectively and with ease, which is crucial for clinical implementation and for long-term market success.

An important distinction between polymer types discussed herein involves the long-term stability of the polymer. Currently, both degradable and non-degradable systems exist. While there are serious limitations on systems that use non-degradable materials, which have limited biomedical applications, polymers of both classes are described in the following section.

A. Non-Degradable Thermosensitive Hydrogels

1. Poly(N-isopropylacrylamide)

One of the most widely studied thermo-responsive polymers is non-degradable poly(N-isopropylacrylamide) (polyNIPAAm), as shown in Figure 1a. PolyNIPAAm has a lower critical solution temperature (LCST) of 32ºC in aqueous conditions, which makes it attractive for biomedical applications.[7] The LCST can also be elevated or reduced to a desired value by incorporating hydrophilic or hydrophobic co-monomers, respectively.[11] Because these co-polymers can be altered to suit the needs of a specific system, the incorporated monomers can also be used to fine-tune additional properties of the material, such as drug-release kinetics, mechanical properties, microstructure, or cellular adhesion.

For example, co-polymerizing polyNIPAAm with acrylic acid results in a polymer that, unlike other hydrogels, does not swell,[12] which could be an exploitable property when incorporating a co-polymer system into the body. Since gelation of this polymer is reversible, this can be advantageous for systems that require replacement, such as the islets in the pancreas.[13] In the pancreas, the cells and polymer can be injected together, but the cells are immobilized upon gelation; when the cells need to be replaced, the gel can be cooled and the polymer/cell solution can be removed.[13] Using the same co-polymer, glucagon-like peptide-1 was also added to the cell/polymer solution to stimulate the islet cells and increase oxygen amounts, which was shown to increase insulin secretions compared to cell delivery without therapeutic agents.[14,15] This suggests that co-polymers are able to perform multiple functions to better improve therapeutic strategies, such as simultaneous cell and drug delivery.

The polyNIPAAm-acrylic acid co-polymer has also been shown to deliver cells and drugs simultaneously, such as chondrocytes and multiple chondrogenic factors.[16] This shows promise in cartilage tissue engineering because these cells are able to differentiate correctly and display well-organized chondrogenesis. Furthermore, when used exclusively for drug delivery, the polyNIPAAm-acrylic acid co-polymer can be combined with poly(ethyl acrylate) as an interpenetrating polymer network to yield an amphiphilic gel that displays pH-induced thermosensitivity at physiological temperatures[17]; this network slows the release of hydrophobic drugs such as daizein, and lacks the burst effect typical of polyNIPAAm. However, the development of polyNIPAAm in biomedical applications may be hindered by the potential toxicity of the NIPAAM monomer,[18] as well as by the non-degradable nature of the polymer.

2. Poloxamer

Another non-degradable system utilizes poloxamer polymers, which also exist under the proprietary name Pluronic®.[14] These polymers are typically triblock, as shown in Figure 1b, and are composed of poly(ethylene oxide) and poly(propylene oxide). Poloxamer 407, a viscous liquid at room temperature that becomes a semisolid transparent gel at body temperature, is of particular interest because it appears to be nontoxic.[14] This poloxamer has been explored as a drug-delivery device for short-term therapies such as pain management,[19] because its release profile does not last more than a few days. However, while low polymer amounts do not alter

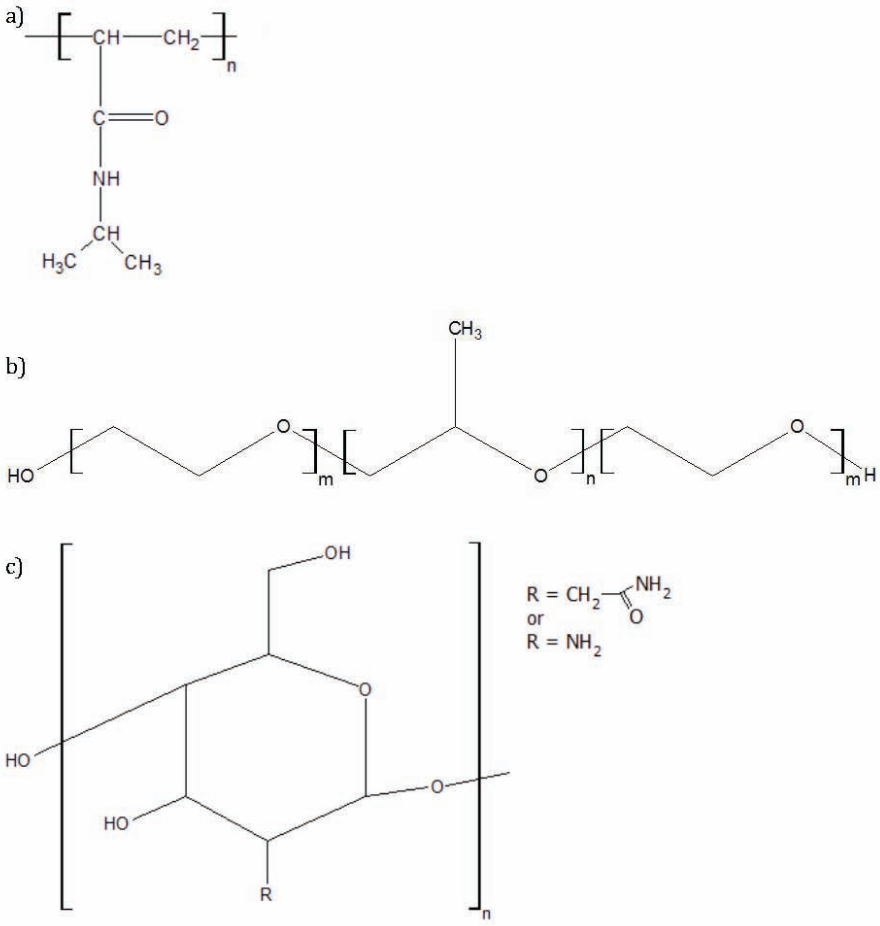


FIGURE 1. Chemical formula of thermoresponsive poly-NIPAAm (a), poloxamer (PEO-PPO-PEO) (b), and chitosan (c).

serum lipids, higher doses have been shown to induce hyperlipidemia in rabbits,[20] which makes the polymer less suitable for repeated dosing, limiting its potential clinical applications. Further, these gels show weak mechanical strength and rapid surface erosion.[7,10,14] Using degradable linkages such as carbonate, ester, and disulfide or urea bonds, some degradability can be achieved,[21] and this can be subsequently tuned by adjusting the length and composition of degradable moieties such as oligo(ester)s, which are inserted into the backbone of the poloxamer polymer.[10,21] Poloxamer solutions have already been administered subcutaneously for the slow release of interleukin-2, urease, and human growth hormone,[22] and their use could extend to other peptides and therapeutic hormones.

B. Degradable Thermosensitive Hydrogels

Given the limitations of non-degradable hydrogels, much interest has recently focused on the creation of degradable, thermo-sensitive hydrogels. Of particular interest are polysaccharides such as chitosan, poly(ethylene glycol) (PEG) block co-polymers, and elastin-related pentapeptides.

1. Polysaccharides

a. Cellulose Derivatives

While most natural polymer solutions form a gel phase when their temperature is lowered, a few polysaccharides display the opposite behavior.[14]

Methylcellulose solutions undergo a phase transition between 40ºC and 50ºC, although this can be lowered through chemical modification.[23] Unfortunately, the non-ideal low viscosity of this polymer can only be overcome through the addition of surfactant,[24] which could pose toxicological problems and may therefore limit clinical development. However, this material has recently had some success as a neural scaffold, because its low protein adsorption and cell adhesion properties are effective for laminin tethering.[25] Methylcellulose has also been combined with hyaluronan to create an injectable hydrogel for the release of neuroprotective and neuroregenerative drugs for 28 days, which is now being further explored in rat models.[26] Metolose®, a non-ionic, water-soluble cellulose ether, has been examined as a transdermal drug-delivery system. In this system, the transition temperature can be modulated with different salts and salt concentrations, which allows for a drug release that is highly dependent on the temperature change.[27]

b. Xyloglucan

Xyloglucan is a polysaccharide derived from the tamarind seed that can form a thermo-responsive gel in water under certain conditions,[14] particularly by the enzymatic removal of galactose.[28] Tamarind seed xyloglucan has three units of xyloglucan oligomers, and when it is partially degraded by beta-galactosidase, it exhibits thermally reversible gelation in dilute aqueous solutions.[14] These gels have been examined for rectal,[29] intraperitoneal,[30] ocular,[31] and percutaneous[32] delivery systems, often showing increased bioavailability and broader concentration profiles of the drugs examined. Specifically, the sustained release of mitomycin C, administered into the peritoneal cavity, and indomethacin, administered rectally, have both shown broader concentration-time profiles and longer residence times with the use of xyloglucan gels.[28]

c. Chitosan

Chitosan, the structure of which is shown in Figure 1c, has been shown to be a promising material for controlled drug delivery[33] and as a cell scaffold.[34,35] Obtained through the alkaline deacetylation of chitin, chitosan is nontoxic, degradable, and biocompatible for many regions in the body.[33] The most promising results to date stem from the technique of graft co-polymerization onto a chitosan backbone, allowing the introduction of functional properties through side chain addition.[33] Chitosan-beta-glycerol phosphate (chitosan-GP) is a typical chitosan used for potential biomedical applications because it has been shown to be biocompatible and has the potential to stimulate mesenchymal stem cell proliferation at certain concentrations.[35] Recently, it was shown that the incorporation of quaternized chitosan, when given in an optimum formulation, can result in 3-minute gel formation at 37ºC.[36] Also, chitosan, while usually present in an acidic formulation of approximately pH 6.2,[37] has been shown to possess temperature-sensitive transition at a neutral pH with the addition of polyol salts.[38] While additives are typically not ideal, this combination has been shown to successfully deliver active growth factors and living cartilage cells in vivo.[38] This polymer can gel at low concentrations of approximately 2% by weight,[37] which furthers its consideration as an injectable material for biomedical applications.

Chitosan hydrogels also show great promise as drug-delivery vehicles. They have demonstrated active paclitaxel release, effectively inhibiting tumor growth and angiogenesis in mice.[39] Also, controlled release of basic fibroblast growth factor-2 (FGF-2) resulted in the induction of angiogenesis in healing-impaired mice,[39] demonstrating the potential for this material to be used to control vascularization. Chitosan and PEG-based in situ gelling materials have also been examined for nasal drug-delivery systems in which the solution is sprayed or dropped into the nasal cavity and subsequently gels to reduce the drug-delivery rate.[40]

In addition to drug-delivery applications, many of these materials have the potential to act as tissue-engineering scaffolds. One popular application is in articular cartilage. This tissue has little capacity for spontaneous healing due to the lack of vascularization. No current repair techniques have successfully regenerated the tissue.[41] However, transplantation of chondrocytes has shown tremendous restorative potential,[41] and therefore a scaffold material that

provides an artificial extracellular matrix for cell adhesion may lead to improved outcomes. Injectable scaffolds represent a noninvasive means to rebuild deteriorated articular cartilage and may provide an adequate matrix for the required support of transplanted chondrocytes. Chitosan-Pluronic hydrogels display thermo-sensitive properties and allow for the proliferation of chondrocytes and the generation of glycoaminoglycan for 28 days.[34] Chitosan-graft-glycolic acid is another injectable scaffold material that has a short gelling time and good chondrocyte viability in vitro.[42] A chitosan-based scaffold co-polymerized with hyaluronic acid has also shown appropriate phase transition and cell survival rates,[43] making it another possible synthetic extracellular matrix material.

Perhaps the most promising material to date is a chitosan-GP system that demonstrates functional matrix deposition and adhesion to both bone and cartilage.[44] This evolved into a hybrid commercial implant named CarGel®, consisting of the chitosan-GP solution and whole blood,[14] which was determined to improve healing.[45] Animal testing of this material has demonstrated accelerated formation of clot scaffolding, specific repair responses including temporary vascularization,[45] and no systemic toxic effects. While this material has shown promise in human clinical studies,[14] evidence indicates that inflammatory responses may occur depending on the injection site.[14,46] Therefore, if successful, this material will likely not be applicable to all potential areas.

Another related application is bone regeneration. While CarGel® may provide some improved bone regeneration,[14] it has not been investigated as rigorously, so other materials may offer improved results. One such material may be an injectable calcium phosphate-chitosan composite scaffold. This material is being investigated as an alternative to the current Food & Drug Administration (FDA)-approved calcium phosphate cement implants, which are brittle and can only be used to reconstruct non-stress-bearing bone.[47] This material showed greater strength and mesenchymal stem cell proliferation and differentiation than the current FDA material, suggesting that it could be used in moderate-load bearing applications while still maintaining new bone growth into the implant.[48] The chitosan-GP gel may also be applicable in bone regeneration, because its injection resulted in decreased macrophage accumulation when rat bone marrow mesenchymal stem cells were entrapped in the material in vivo.[49] From these results, it is clear that an injectable, chitosan-based scaffold could result in a bone-regeneration scaffold in the near future.

2. PEG

PEG, a hydrophilic and compatible material, can be combined with hydrophobic, degradable polyesters to yield thermo-responsive hydrogel systems. PEG can be incorporated with poly(D,L-lactic acid-co-glycolic acid) (PLGA) as a triblock co-polymer to give a transparent gelling material with a transition temperature of 37ºC and a critical gel concentration of 16% by weight.[14,50,51] This gelling behavior can be manipulated by altering the concentration, as shown in Figure 2, and the monomer feed ratios. Because it maintains its three-dimensional structure, PEG has been shown to gel quickly[52] and form a strong implant that induces little tissue irritation. While only low-molecular-weight PEG can be removed by the kidneys,[53] it can be resorbed from the injection site[14] and therefore the co-polymers are considered degradable.[10] To overcome molecular-weight limitations regarding PEG length and degradability, short-chain graft co-polymers are frequently employed,[14] as discussed below.

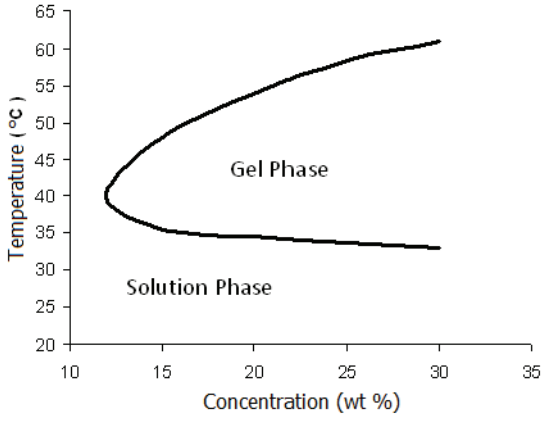


FIGURE 2. Phase diagram of PLGA-PEG-PLGA triblock co-polymer (adapted from Yu and Ding,[10] with permission of The Royal Society of Chemistry).

a. PLGA

When PEG-g-PLGA and PLGA-g-PEG are combined, the degradation of the material can be manipulated to occur between 1 week and 3 months, while showing in vivo cartilage repair.[10,14] Sustained release of a drug grafted with PEG from the PLGA-PEG-PLGA co-polymer has also been explored; the release is believed to be diffusion-controlled initially, but controlled by both diffusion and the hydrogel erosion in later stages.[54] This release technique has shown some success in mice,[54] further demonstrating its potential efficacy.

Many of these thermosensitive hydrogels display sustained protein-release profiles,[55] which is a current challenge of many existing noninvasive drug-delivery systems. Additionally, the prolonged effect is localized to the injection site, providing therapeutic concentrations of pharmaceuticals within target tissues, which can be useful for treatments such as cancer therapy, while potentially reducing side effects caused by the cytotoxic drugs. A particular example of this is the PLGA-PEG-PLGA co-polymer system ReGel®, which has been used to release the drug paclitaxel. The in vitro experiments showed release over 50 days, compared with 1 day for the existing poloxamer 407 method.[7,14,56] In vivo, the material degraded in 4 to 6 weeks, showed slow clearance from the injection site, and had little distribution into other organs.[56] When used against human breast cancer tumor xenographs, this combined product, called OncoGel®, showed superior results and fewer side effects than the currently available Taxol® drug product.[56] OncoGel® has entered Phase II clinical studies and the ReGel® system is also being explored for insulin delivery,[7,14] with the potential to achieve basal insulin levels in 1 week with a single injection.[7]

b. Polycaprolactone

Like PLGA, polycaprolactone (PCL) is a degradable, thermo-responsive polymer with a concentration-dependant phase transition temperature.[57] PCL offers an attractive alternative to PLGA, because PCL-PEG multiblock co-polymers exhibit solution stability, a wider gel window, and a higher gel modulus.[10] Additionally, these co-polymers have a powder consistency at room temperature and are easier to handle than sticky PLGA pastes.[10] Furthermore, the degradation product caprolactone results in weaker acidification than lactic acid or glycolic acid, which can help prevent non-bacterial inflammation in vivo.[10] Further tests have shown continuous drug release and cell differentiation,[10,57] suggesting that the material has potential in biomedical applications. Interestingly, poly(ε-caprolactone-co-D,L-lactic acid), when combined as a block polymer with PEG and sulfamethazine oligomers, can result in a dual-response gel that requires certain pH and temperature conditions to gel. This could improve clinical applications further, because it would prevent gelation during syringe injection and make clinical handling and processing simpler.

3. Elastin-like Peptides

Elastin-like peptides are artificial polypeptides that have the conserved sequence Val-Pro-Gly-Xaa-Gly. Because they are monodisperse, temperature responsive, and biocompatible with many systems,[59] they have great biomedical potential. These materials have been combined with collagen-derived tripeptides to yield a cell-adhesive amphiphilic co-polymer that shows a temperature-induced phase transition. The phase transition is believed to be the result of a conformational change in the polypeptides.[10] This is currently used as a more efficient method for protein purification, though its efficiency, specificity, and reversibility could yield many potential applications. While these materials may show great promise, they are highly concentration dependent, which may limit their applicability. Also, they have received relatively little study in the biomaterials and drug-delivery field.

C. Summary of Thermo-sensitive Materials

Injectable, in situ forming hydrogels have great potential for use in drug-delivery and tissue-engineering applications, because they can improve delivery efficacy through minimally invasive procedures. In particular, thermo-sensitive materials hold particular promise in that they do not require the use of solvents

or cross-linkers, which can result in toxicity or damage to the entrapped materials. Degradable materials have the potential advantages of a lowered inflammatory response and more complete healing. Further, the degradation rate can be controlled to yield better release profiles and cell-migration patterns.

Given the simplicity of mixing cells or drugs into the materials, the general compatibility of many of the materials, the potential for controlled release profiles, and the ability to deliver therapeutic agents to hard-to-access regions of the body, thermo-responsive materials offer exciting potential for use in a vast array of biomedical applications.

III. PHOTO-RESPONSIVE MATERIALS

While photo-active groups have been used extensively for many decades in the generation of novel polymers, photo-responsive materials have only recently been introduced to the biomaterials literature. Photo-responsive properties may be introduced into polymers through the incorporation of different photo-sensitive molecules that undergo chemical and physical changes in response to specific types of light. Photo-responsive systems have fast response times because they do not depend on diffusion. Furthermore, the application of laser and light has been widely established in medicine and can lead to high spatial and intensity control.[5,61] The most common application of light stimuli in polymer chemistry is light-activated polymerization. Photo-initiators absorb light of 250 to 450 nm to form reactive intermediates in the form of free radicals, cations, or anions that initiate polymerization.[62] The fast curing rates and ambient temperature requirements of photo-polymerization make it gentle on biological materials and highly controllable.[63] Photo-polymerization has been widely applied to generate polymers in numerous applications. However, the focus in this review will be on systems that respond to light post-synthesis. These systems represent a lower risk to biological systems because they typically do not require the use of external photo-sensitizers and therefore minimize the potential for leaching of undesired and potentially toxic side products.[5]

Photo-responsive systems are beneficial for on-demand alterations of materials in vitro or for the alteration of scaffold and drug-delivery biomaterial properties in vivo. More specifically, ophthalmology, gynecology, dermatology, otolaryngology, gastroenterology, and physiotherapy could benefit greatly from the application of these smart photo-responsive biomaterials because they have direct applications in medical areas that are already well-versed in the application of lasers.[61]

Research into photo-responsive polymers has focused on several biomaterial applications. Cross-linked gels with alterable properties can act as cell scaffolds and drug-delivery systems. Micelles and liposomes, which are monolayer and bilayer phase-separated spheres formed from amphiphilic molecules, as well as other nano- and microparticles, are of interest in drug delivery. Reversible and alterable surface properties have also been investigated for cell growth and protein patterning.

Light-activated molecules bound to polymers can create reversible and irreversible systems to control drug delivery, cell scaffold architectures, surface patterns, and nanoparticulates. Photo-isomerization, photo-dimerization, photo-degradation, and photo-triggered mechanisms will be elaborated upon to illustrate recent photo-responsive polymer research that shows potential in biomaterial applications.

A. Reversible Systems

1. Photo-isomerization

When bound to polymers, molecules that undergo photo-isomerization between cis and trans forms with different wavelengths of light can introduce photo-reversible properties. Polymers with azobenzene,[64] spirobenzopyran, and variations thereof[65] incorporated as functional side chains undergo changes in hydrophilicity/hydrophobicity or structural changes as a function of light exposure, as illustrated in Figure 3.[66] Polymer systems containing photo-isomerizing groups are currently the most investigated polymer-based, reversible photo-responsive materials.

Photo-isomerizing groups can be used to modify the rheological and cross-linking properties of polymer gels to create materials that may be suitable for drug delivery and in situ gelation. For example, azobenzene has been shown to alter the rheologi-

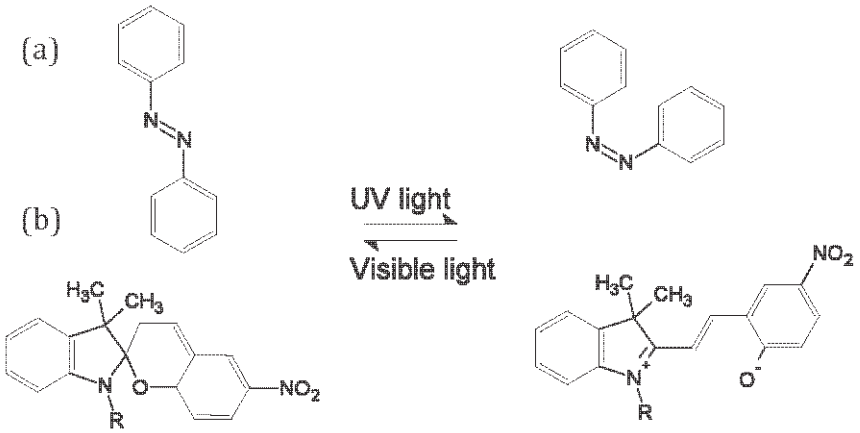


FIGURE 3. Azobenzene (a) and spirobenzopyran (b) isomerization occurs with UV and visible light.

cal properties of polymers by creating hydrophobic domains within polymer networks that bridge reversibly with UV light to cause cross-linking.[67] The length decrease from 9.0 to 5.5 angstroms associated with the *trans* to *cis* isomerization of azobenzene[68] can be used to alter the cross-linking mesh sizes of polymer gels when azobenzene is grafted to multiple polymer chains.[69,70] Release of poly(styrene sulfonic acid) from these gels showed slightly modified drug delivery with UV treatments. However, the investigators acknowledged that this was not to a great extent and was likely due to the small, dimensional changes between the *trans* and *cis* isomers of azobenzene.[69] Cyclodextrin-containing polymers and azobenzene-modified polyacrylate can form clustered complexes in solution that cause an increase in viscosity. Complexation occurs when azobenzene is a *trans* isomer, and therefore light can be used as a photo switch to increase or decrease the amount of *trans* isomers, and subsequently the complexes that form, to control the viscosity of the solution.[71] While photo-isomerizing molecules clearly introduce reversible changes into bulk polymer systems, attempts at altering cross-linking density and drug delivery have shown minimal influences on changing the diffusion and delivery rates.[69,72]

Micelles composed of amphiphilic diblock co-polymers that can undergo dissolution and reformation with light stimuli can be created by incorporating photo-isomerizing groups as a side chain on small diblock co-polymers.[73] If *trans* isomers of azobenzene are bound to one block of a hydrophilic block co-polymer, it will become hydrophobic and drive the formation of micelles. Photo-isomerization with UV light will change the azobenzene to its *cis* form, making the entire diblock co-polymer hydrophilic and causing the micelle to dissolve. For example, co-polymers of hydrophilic poly(*tert*-butyl acrylate-co-acrylic acid) with azobenzene-bound polymethacrylic groups undergo reversible micelle formation. As illustrated in Figure 4, light alters the amphiphilicity of the block co-polymers with the *cis/trans* transformation of azobenzene with UV and visible light, thereby causing the micelles to form and dissociate with light treatment.[73-74] Liposomes, which are similar to micelles but contain a sphere formed by a bilayer, can be created with phosphatidyl choline that contains azobenzene. Upon UV irradiation, kinks form in the phosphatidyl choline from azobenzene isomerization, which increases the permeability of the liposomes, allowing contents such as the anticancer drug doxorubicin to release.[75] In another system, UV irradiation of core-cross-linked micelles containing azobenzene allows *cis* to *trans* isomerization to reversibly produce hollow spheres.[76] Amphiphilic azobenzene-containing macromolecules can form vesicles that can aggregate after treatment with UV light[77] or can form bilayers with alterable surface characteristics.[78]

Nanogels and colloidal spheres may be produced by incorporating photo-isomerizing molecules into the backbone of the polymer. Dextran nanogels with

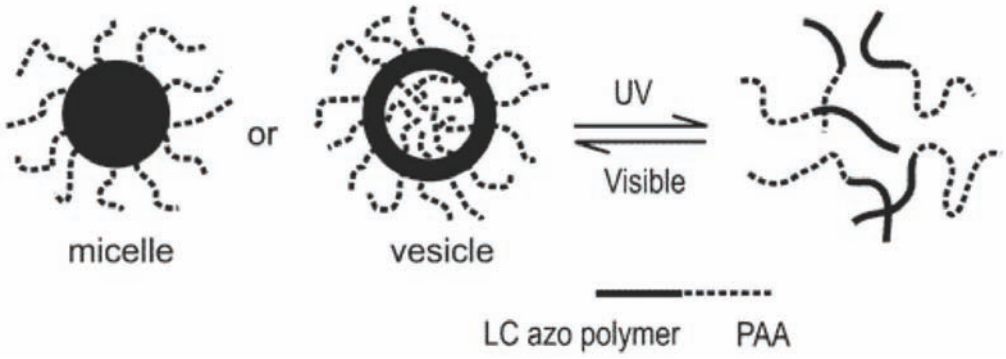


FIGURE 4. When diblock co-polymers are changed from having a hydrophilic and hydrophobic end to being completely hydrophilic, they dissolve and release their contents. (Reprinted with permission from Wang et al.[74] Copyright 2004, American Chemical Society.)

multiple azobenzene groups bound along the backbone have demonstrated changes in size as a function of light, because hydrophobic isomers increase aggregation and physical cross-linking, effectively decreasing nanogel size.[72]

The conformation of polymers in solution may also be controlled via grafting of azobenzene to potentially create sensory devices, shape-memory materials, or molecules that can self-assemble. When azobenzene is incorporated as a side group along polymer backbones, polymers in solution can take a helical shape that may be reversibly flipped from a left helix to a right helix or irreversibly destroyed.[79] Azobenzene-modified polypeptides, polyaramides, polyesters,[80] and poly(ester amides) have shown reversible helical shape responses.[81-82] Such effects may well suit the area of photo-controlled self-assembly, which can create photo-switchable architectures and assemblies.[79] Surfactants are used to create colloids and emulsions. The association between a neutral surfactant with azobenzene-containing polyacrylic acids was found to weaken after UV light treatment, which turned azobenzene into its polar cis form, causing disruption of the system. This system may be applicable to micelle and interface applications.[83] Colloidal particles containing azobenzene groups within the polymer backbone deform with light treatments,[84] and emulsions have also been reversibly directed from oil droplets in a water phase and from water droplets in an oil phase depending on the isomer form of azobenzene in the backbone of poly(sodium acrylate).[85]

Unique systems have been investigated that use photo-induced patterning and alterable protein affinity. Photo-responsive spirobenzopyran materials have been demonstrated to create cell micropatterning[86] and microvalves for on-chip fluid control.[87] As shown in Figure 5, molecularly imprinted materials with shaped pores containing azobenzene have reversible uptake and release of caffeine and other pharmaceuticals due to *cis/trans* photo-isomerization.[88,89] Interestingly, peptides have an affinity for the *cis* but not the *trans* isomer form of azobenzene, a fact that may be used to isolate peptides from solution.[90] Binding of azobenzene-capped polymers to streptavadin may be used to control its binding affinity with the *cis/trans* conversion, causing either the collapse of the polymer onto the surface to block activity or the extension of the polymer away from the surface to unblock the active site and allow activity.[91] Photo-bending polymer films based on azobenzene have also been created, with UV light inducing bending and visible light flattening the films.[92,93] Surface patterning of azobenzene polymer films with wave-like patterns can be obtained using sinusoidal light interference pattern exposures,[93,94] and wettability of azobenzene-modified surfaces can also be altered reversibly with light treatments.[95]

Photo-isomerization is a common method of introducing photo-reversible changes in polymer systems. Although in some cases, the photo-induced changes are subtle, the isomerizing groups are highly reversible and require no photo-sensitizers, and these are important benefits in creating photo-reversible biomaterials.

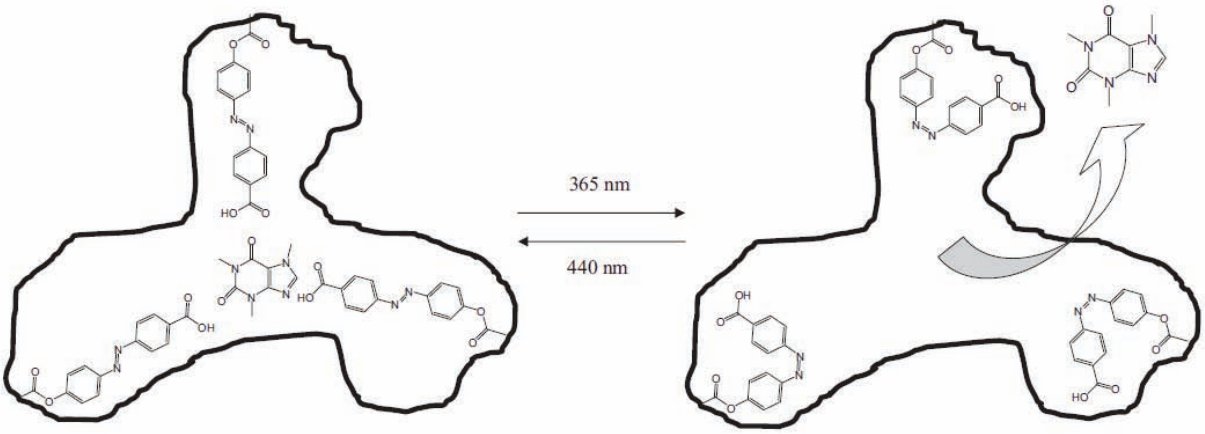


FIGURE 5. The uptake and release of caffeine from molecular imprints controlled by azobenzene and light treatments. (Reprinted with permission from Gong et al.[90] Copyright Wiley-VCH Verlag GmbH & Co.)

2. Photo-responsive Dimerization

Molecules that have the ability to reversibly dimerize with a second molecule can be used to form reversible, covalent modifications to polymer systems. By modifying polymers with photo-dimerizing molecules, photo-cross-linking/binding gels, micelles and surfaces can be designed with minimum side-product formation and fast curing.[96] As illustrated in Figure 6, anthracene,[97] cinnamylidene acetate, nitrocinnamate, coumarin,[98] furan,[99-100] chalcone,[101,102] and maleimide[103,104] derivatives can photo-dimerize. The first four molecules are particularly widely studied in the development of photo-reversible synthetic and natural polymers. Nitrocinnamate, cinnamylidene acetate, and coumarin undergo isomerization followed by reversible [2π+2π] photo-addition to form cyclobutane rings,[66,98,105,106] while anthracene undergoes reversible [4π +4π] photo-cycloaddition with dimerization at wavelengths over 300 nm and dissociation at wavelengths less than 300 nm.[97,107,108] When covalently incorporated into polymer systems, these molecules can dimerize and dissociate with specific wavelengths, thus altering bulk polymer properties and in some cases causing changes that may be reversed with UV wavelengths under 300 nm.

Photo-cross-linking of hydrophilic polymers with photo-dimerizing groups can create gels that bind proteins, support cell growth, or alter the diffusion of proteins. PEG-based hydrogels with photo-dimerizing groups have been widely investigated in recent years due to the compatibility of PEG in biological systems.[109] A cinnamylidene acetate-terminated four-arm PEG has been demonstrated to form gels with UV light >300 nm and to have reversible swelling ratios with alternating >300 nm/254 nm UV light exposures.[110] Protein flux through membranes of this material could be lowered with >300 nm light treatment,[111] and active coumarin-bound organophosphorus hydrolase could be photo-immobilized onto the gels, creating enzymatically active hydrogel materials.[112] However, sol/gel transition does not occur with these materials due to a competing dimerization that also occurs at low UV wavelengths of 254 nm, which causes cross-linking to maintain the gel.[113] Nitrocinnamate-terminated eight-arm PEG can photo-cross-link to form gels with reversible surface and mechanical properties, as illustrated in Figure 7.[114,115] These gels have been demonstrated to slow the release of FGF; 70% was delivered over a 5-day period with a lowering to 50% with UV treatments.[116] Anthracene-terminated eight-arm PEG that undergoes photo-cross-linking has also been synthesized and demonstrated light-responsive, reversible surface roughness.[96] Photo-dimerizing groups capped on four- and eight-arm PEG formed photo-sensitive systems, but these were not completely reversible in many cases. However, reversible, cyclic polymer chain growth of coumarin-capped PEG chains with

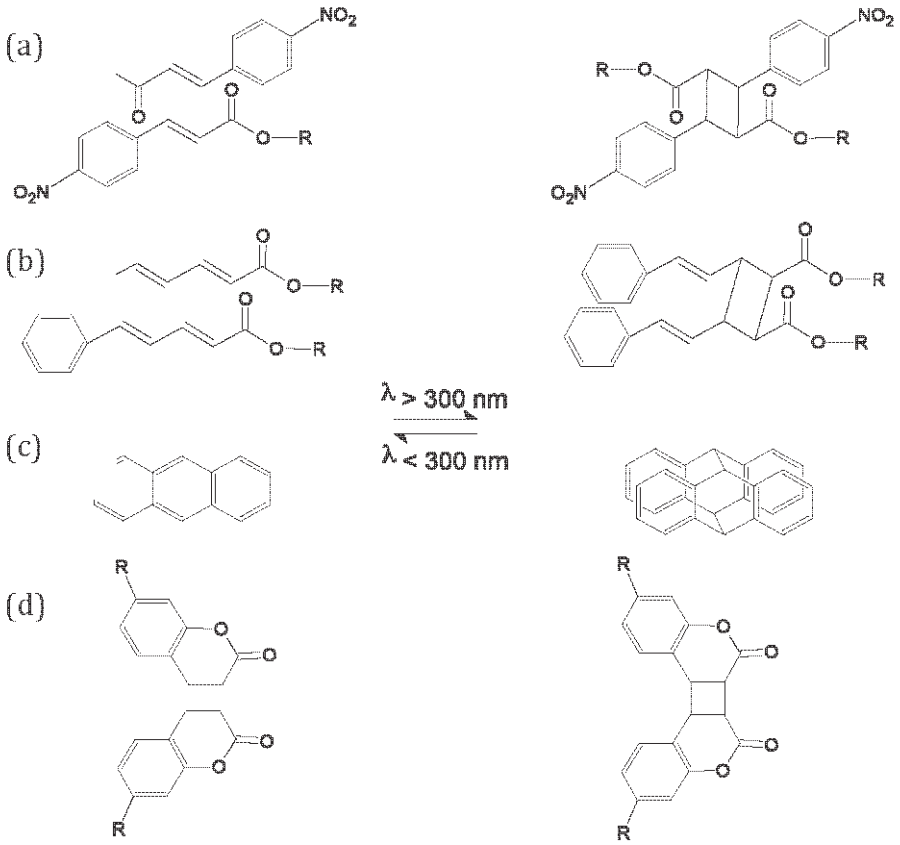


FIGURE 6. Nitrocinnamate (a), cinnamylidene acetate (b), anthracene (c), and coumarin dimerization and dissociation/de-dimerization with specific UV wavelengths.

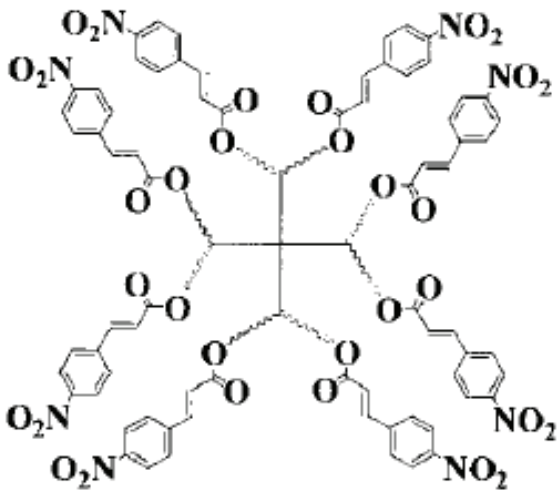


FIGURE 7. Eight-arm polyethylene glycol nitrocinnamate used to create photo-sensitive PEG-based hydrogels. (Reprinted with permission from Zheng et al.[117] Copyright Wiley-VCH Verlag GmbH & Co. KGaA)

alternating UV treatments has been demonstrated in solvent cast films.[117]

Natural hydrogels bound with photo-dimerizing groups have the potential to create compatible biomaterials with the potential to undergo hydrolytic or biodegradation. Gelatin with nitrocinnamate groups incorporated directly onto its backbone has the ability to photo-cross-link with 365-nm and to have topographical changes with 254-nm light treatment.[118] Anthracene-bound dextran creates a photo-sensitizer with hydrophobic domains.[119] Amino-acid-based organogels have been created that undergo gelation with 365-nm light treatments.[120] Other polymers have been created via photo-dimerization but may degrade via other means. For example, polyoxazoline gels have been created that cross-link via the dimerization of anthracene in their backbone and solubilize via disulfide bond cleavage from bonds adjacent to the anthracene groups.[121]

While many hydrophobic systems have been created with photo-dimerizing molecules, only those with direct potential in biomaterial applications such as surface patterning will be discussed. Hydrophobic polymers seem well-suited for hydrophobic anthracene. For example, block co-polymers of anthracene-terminated polystyrene and anthracene-terminated polymethylmethacrylate can be dimerized to bind the two different chains[122] and control surface patterns of films.[123] Polydimethylacrylamide co-polymers that contain coumarin side chains have been used to attach and photo-immobilize biopolymers such as heparin and then to control their release.[124] Surface treatments can be used to create photo-induced honeycomb patterns that disappear and reappear with light treatments to create a reversible lithographic technique from anthracene-containing poly(methylmethacrylate).[125] Poly(ε-caprolactone)-based polymers undergo hydrolysis to degrade in vitro and in vivo. Coumarin has been introduced into poly(ε-caprolactone-*co*-trimethylene carbonate) to photo-cross-link the polymer and create surface-erodible matrices.[126,127]

Shape-memory properties allow polymer films and gels to form stimuli induced shapes such as coils. Cinnamylidene-based poly(hydroxyethylmethacrylate) can be photo-cross-linked into a shape that relaxes upon de-cross-linking at UV light treatment under 260 nm and reforms upon treatment above 630 nm.[128] Photo-reversibly induced film-bending, as illustrated in Figure 8, is another interesting phenomenon that was demonstrated with coumarin-containing supramolecular polymers, and has been attributed to cross-linking on one side of the film.[129]

Nanoparticles, microgels, and micelles have also been synthesized via photo-dimerization to create rapidly responding systems that may cluster or have triggered release of drugs. Photo-dimerization can also be used to reinforce micelles or nanoparticles that may dissolve to release drugs or other substances. Diblock polymers of PEG and poly[2-(2-methoxyethoxy)ethyl methacrylate-co-4-methyl-[7-(methacryloyl)oxyethyloxy] coumarin] and diblock polymers of PEG and poly(coumarin methacrylate) can form micellar aggregates that could be photo-reversibly core-cross-linked with light >300 nm and had a reduction in cross-linking with light <260 nm, as illustrated in Figure 9.[130,131] Polyacrylic acid can be

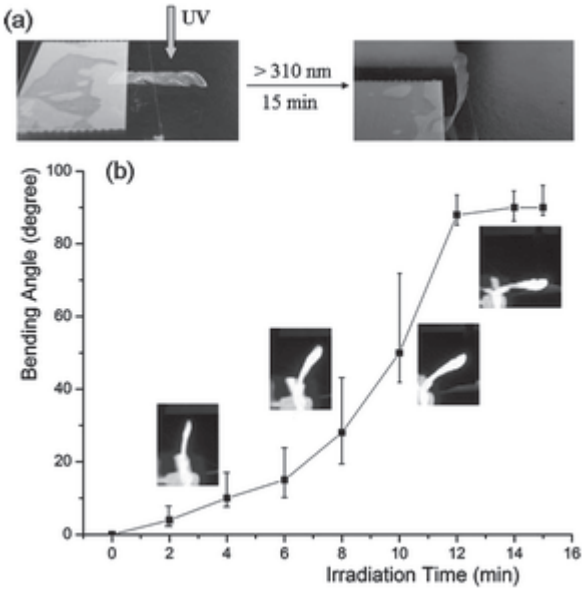


FIGURE 8. Films of poly(4-vinyl pyridine) that contain 7-(carboxymethoxy)-4-methylcoumarin have UV light-induced bending. (Reprinted from He et al.[132] Reproduced with permission of The Royal Society of Chemistry.)

used to assemble similar nanoparticles into a film to create photo-reversible pore sizes.[132] Nitrocinnamate has been used to create "micro-containers" from emulsions with photo-dimerization creating cross-linked shells that may be solubilized with 254-nm light.[133] Cinnamate-based hyperbranched polymers have also demonstrated an ability to create cross-linked nanoparticles with controlled hydrolytic degradation.[134] Loading and release of quantum dots has been demonstrated with microgels of nitrocinnamate-containing PEG or gelatin created by photo-cross-linking within inverse micelles.[135] Light-induced clustering and de-clustering of nanoparticles can be induced in coumarin- and nitrocinnamate-containing poly(organosiloxane) nanoparticles to control colloidal dispersions.[136]

Photo-dimerizing molecules create polymer systems that may be effectively cross-linked with UV light. Dimerization can cause gelation and reinforce micelles and nanoparticles. While gel-sol transitions are not always possible due to competing dimerization that occurs during dissociation at low UV wavelengths (<300 nm), effective changes can still be induced with UV stimuli. However, par-

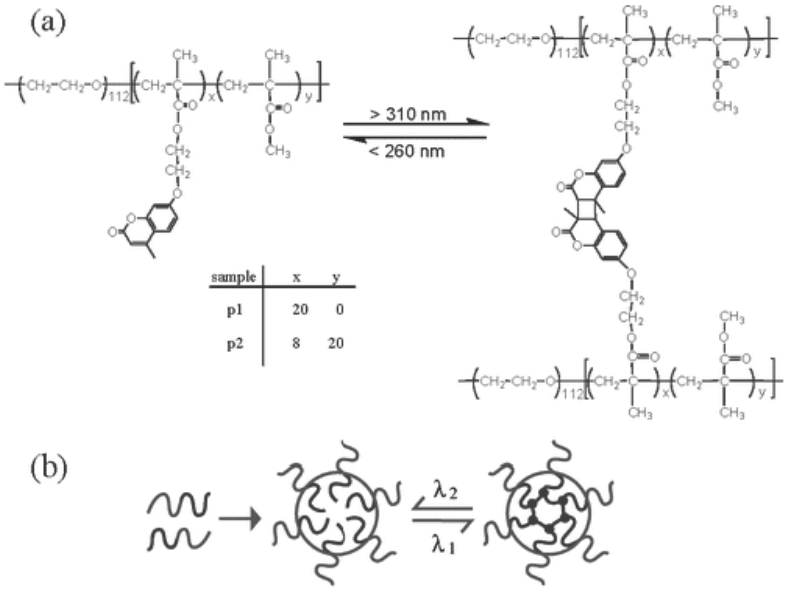


FIGURE 9. The structure of micelles that can undergo core cross-linking with UV light >310 nm and dissociate with UV light <260 nm. (Reprinted with permission from Jiang et al.[134] Copyright 2007, American Chemical Society.)

tial de-dimerization can effectively cause property changes in surfaces and micelles.

B. Irreversible Systems

1. Photo-responsive Degradation

Photo-induced degradation of polymer systems would be beneficial for applications of biomaterials that require property alterations post-synthesis or that need to be removed after their use/life. While photo-degradation is an irreversible process, it can be used to alter the properties of polymer biomaterials after synthesis.

External agents may be introduced into a system to initiate degradation upon UV exposure. One such agent is riboflavin, which has been shown to degrade polymers such as alginate[137,138] and hyaluronic acid[139,140] once activated with UV light. Alginate gels containing riboflavin have efficient photo-sensitized degradation of the alginate chains via oxidative cleavage of glycosidic bonds.[137,138] Photo-activatable external agents may also be added to self-assembled polymers to induce depolymerization to break down polymers.[141] While these systems offer effective methods to photo-degrade polymers, these external agents may leach and therefore may not be suitable for some biomaterial systems.

Covalent modification of polymers with photo-labile groups has the potential to alter properties in incremental fashions. Nitrobenzyl-based groups are photo-cleavable and well used in these applications to create scaffolds and patterned surfaces and to create dissolvable nanoparticles.[142,143]

Cross-linking of a system with photo-labile molecules will allow the de-cross-linking of the matrix or cleavage of grafted molecules to alter polymer properties in an incremental fashion. For example, nitrobenzyloxycarbonyl-linked star polymers can degrade with UV light.[144] Interestingly, materials may be synthesized with subsequent light exposure used to selectively cleave away proteins and other molecules, resulting in the creation of patterned scaffolds and surfaces that have the ability to direct cell growth.[142] PEG hydrogels containing UV-cleavable nitrobenzyl ether-derived moieties have been used to create channels to control cell migration and to release bound peptides to control cell growth.[145] Photo patterning has also been achieved, for example, with nitroveratryloxycarbonyl-bound

groups that could be "photo-removed" from surfaces.[145] Polymer brushes cleaved with UV light can alter surface properties. Methacrylate-based brushes end-capped with hydrophilic groups via photo-labile nitrobenzyl connections become hydrophobic with UV treatments that cleave the nitrobenzyl groups, exposing the hydrophobic methacrylate.[146]

Photo-degradable micelles, nanoparticles, and liposomes can create triggered drug release systems. For example, diblock hydrophilic co-polymers can have one end altered to become hydrophobic with the binding of hydrophobic dyes via photo-cleavable pyrenylmethyl esters. This creates amphiphilic diblock co-polymers that form micelles. With UV exposure the dyes are cleaved off and the entire diblock co-polymer becomes hydrophilic, thereby allowing the micelles to dissolve and release their contents.[143,147] Liposomes with photo-induced changes in permeability can promote encapsulated drug release and have been created with photo-reactive lipids such as 1,2-bis(10-[2',4'-hexadienoyloxy-decanonyl]-*sn*-glycero-3-phosphocholine). UV light exposure causes discontinuities in the liposomes, increasing their permeability.[148,149] Photo-labile lipids that degrade to destabilize the lipid bilayer of liposomes may also promote the release of molecules upon UV light exposure, as has been demonstrated with the entrapment and release of calcein dye.[150] Closed-shell nanoparticles that incorporate nitrobenzyl groups have been shown to cleave with UV light to release their contents.[151]

2. Photo-triggered Systems

The use of UV light to "photo-trigger" the release of active molecules is a novel approach for drug delivery. One clever example involves the simultaneous activation and release of drug conjugates from scaffolds with UV light. Light-sensitive conjugates of acetyl salicylic acid, ibuprofen, and ketoprofen were created and incorporated into methacrylate-based scaffolds. Upon UV exposure, they were cleaved into their active form and diffused out of the scaffold.[152] Photo-triggered release has also been applied to the fragrance industry, with photo-polymers showing long photo-induced release times of fragrant aldehydes and ketones with 350-nm light to create longer-lasting perfumes.[153]

Photo-lysis may be used to alter emulsion properties. The destabilization of a microemulsion can be induced by the photo-degradation of a photo-destructible surfactant that causes an initial increase in droplet size followed by separation of the oil and water phases, which can be potentially used as a system to release molecules from microemulsions.[154] Similarly, light irradiation can trigger globular hydrogel formation from large vesicles due to changes in hydrophilicity caused by UV irradiation of 2-diazo-1,2-naphthoquinone-containing co-polymers.[155] This globular formation can trigger the release of encapsulated dye molecules.[155] As demonstrated with chlorambucil, binding of one side of dimerized coumarin to a drug molecule and the other side to a polymer can create a photo-triggerable release system, with UV-induced de-dimerization releasing the drug from the polymer system.[156]

Photo-triggered release can be used to both activate drugs and change the properties of colloidal systems to release molecules. These irreversible systems are particularly suited for inducing drug release.

C. Photo-responsive Polymers that induce Secondary Reactions

Photo-activated molecules have the potential to influence the properties of surrounding polymers to create a secondary reaction. For example, light may cause photo-cleavage of a molecule into charged groups. Hydrogels containing triphenylmethane leuco-derivatives dissociate under UV light, introducing ionic groups into the gels that then cause an increase of osmotic pressure within the hydrogel.[5,157-158] Photo-induced creation of other molecules may cause changes in the system. For example, microcapsules containing azobisisobutyronitrile produce nitrogen gas under light exposure, which induces the eventual rupture of the capsules and the release of their contents.[6] Light-induced temperature changes may also cause secondary reactions in thermo-responsive polymer systems, which is discussed in a later section. It is important to realize that while these systems function well, they have slower response times relative to strictly photo-responsive systems because diffusion must also occur through the polymer system.

D. Summary of Light-Responsive Polymer Systems

Light-responsive polymer systems have the potential to create drug-delivery, cell-scaffold, surface-patterned, and nanoparticulate biomaterials with quick response times and low side-product formation. Photo-isomerization and photo-dimerization molecules introduce potentially reversible changes in polymer properties, and photo-degradation and photo-triggering mechanisms can cause timed irreversible changes in polymer properties after synthesis. Photo-responsive systems offer the potential to create unique systems that are easily altered with light or laser and may be tailored toward controlled disease treatment and tissue regeneration.

IV. THERMO- AND PHOTO-RESPONSIVE POLYMER SYSTEMS

A classic example of polymer systems that respond to both thermal and light stimuli are those that incorporate photo-sensitive molecules that produce heat upon photo-irradiation, which leads to phase transitions in thermo-sensitive polymers. Trisodium salt of copper chlorophyllin produces heat upon irradiation of visible light, and this has been shown to promote phase transitions of thermo-responsive polymers such as polyNIPAAm.[159] Gold nanorods have also shown the ability to photo-thermally induce volume changes in polyNIPAAm-acrylic acid microgels with near-infrared light,[160] and gold nanoparticles have demonstrated the ability to actuate the LCST of polyNIPAAm-co-acrylamide microgels with light.[161]

Photo-isomerizing groups can also be used to alter the LCST of thermo-sensitive polymers in vitro or in vivo. Research shows that the hydrophobicity of incorporated molecules can alter the LCST of polyNIPAAm with increased hydrophobicity, thus lowering the LCST. Azobenzene end group-functionalized polyNIPAAm exhibits different phase-transition temperatures dependent on the ratio of *cis* to *trans* isomers of azobenzene, with hydrophobic *trans* isomers promoting lower LCSTs with differences up to 10ºC.[162] Similarly, polyNIPAAm with azobenzene-containing acrylamides exhibits the same trend of lower LCSTs after visible light inducing the formation of *trans* isomers and higher LCSTs after UV light inducing the formation of *cis* isomers.[163] Surface modification of silicon with grafting of azobenzene-containing poly(N,N-dimethylacrylamide)-based polymers results in changes to the wettability because LCST is modified in a similar fashion with UV and visible light.[95] Spiropyran- and spirobenzopyran-based polyNIPAAm microgels also had different LCSTs depending on UV treatment and the resulting hydrophobicity from the cis versus *trans* isomerization.[164,165]

Hydroxypropyl methylcellulose (HPMC) undergoes sol-gel transitions at high temperatures. The gelation temperature can be altered by modifying its hydrophobicity with photo-isomerizing groups and external additives. Cyclodextrin can complex with the *trans* form of azobenzene, changing it from hydrophobic to hydrophilic. Therefore, complex formation between cyclodextrin and HPMC-azobenzene polymers can be controlled with light treatments, as illustrated in Figure 10.[166-167] The complex formation will have an overall effect on the hydrophobicity and therefore gelation temperature of the HPMC. After visible light treatment, cyclodextrin interacts with the *trans* form of azobenzene, which reduces

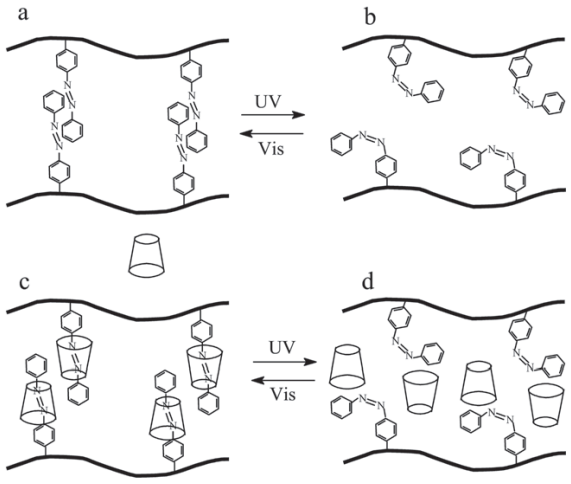


FIGURE 10. Azobenzene grafted to HPMC can undergo isomerization with UV light (a,b). Cyclodextrin can complex with the trans isomer of azobenzene but not the cis isomer (c,d). Reprinted with permission from Zheng et al.[172] Copyright Wiley-VCH Verlag GmbH & Co.)

the hydrophobicity of the trans groups, resulting in a large increase in the gelation temperature. After UV treatments to transition azobenzene into its *cis* form, the gelation temperature of HPMC-azobenzene with cyclodextrin decreases.[167] Gelation temperature can also be altered with cyclodextrin added to azobenzene-modified polyNIPAAm, diethylacrylamide, and dimethylacrylamide.[168]

In some studies, photo-dimerizing groups with thermo-responsive polymers have been combined to create micro- and nanoparticulates or to cause aggregation. Heat has also been used to create micelles, which followed by irradiation with UV light to reversibly core cross-link them via coumarin dimerization.[130] In addition, when coumarin is incorporated into the backbone of poly(ether amine) also containing short poly(ethylene oxide) chains, light- and temperature-sensitive nanoparticles that may be clustered upon irradiation with UV light are created.[169] Thermo-driven shape memory can be maintained in coumarin-cross-linked poly(ε-caprolactone) polymers that undergo reversible cross-linking and de-cross-linking with light.[170]

V. CONCLUSIONS

Stimuli-responsive materials have received significant attention in recent years for use in biomedical applications due to their ability to respond to real-time environmental stimuli and undergo controllable, preprogrammed responses. The predictable nature of stimuli-induced property changes is allowing ever-increasing opportunities to tailor materials for applications such as drug delivery, cell encapsulation, and tissue engineering, allowing controllable release profiles, minimally invasive delivery, and intelligent synthetic extracellular matrix analogs. Thermo-responsive and light-sensitive materials are two families of intelligent materials that have demonstrated tremendous potential for biomedical applications and are rapidly evolving to suit the demands of various systems. Through co-polymerization, polymer grafting, and the development of interpenetrating networks, it is possible to bring together multiple stimuli-responsive materials to yield elegant combinations capable of responding to a multitude of parameters. As the body of literature in these fields expands, so too will the seemingly limitless number of possible combinations, ultimately allowing biomaterials to be designed and optimized specifically for their desired application.

REFERENCES

1. Bawa P, Pillay V, Choonara YE, du Toit LC. Stimuli-responsive polymers and their applications in drug delivery. Biomed Mater. 2009;4:1–15.
2. Schmaljohann D. Thermo- and pH-responsive polymers in drug delivery. Adv Drug Deliv Rev. 2006;58:1655–70.
3. Andresen TL, Jensen SS, Jørgensen K. Advanced strategies in liposomal cancer therapy: Problems and prospects of active and tumor specific drug release. Prog Lipid Res. 2005. 1;44(1):68–97.
4. Smolensky MH, Peppas NA. Chronobilogy, drug delivery, and chronotherapeutics. Adv Drug Deliv Rev. 2007;59:828–51.
5. Qiu Y, Park K. Environment-sensitive hydrogels for drug delivery. Adv Drug Deliv Rev. 2001;53:321–39.
6. Kost J, Langer R. Responsive polymeric delivery systems. Adv Drug Deliv Rev. 2001;46:125–48.
7. Klouda L, Mikos A. Thermoresponsive hydrogels in biomedical applications. Eur J Pharma Biopharma. 2008;65:34–45.
8. Ishihara M, Fujita M, Obara K, Hattori H, Nakamura S, Nambu M, Kiyosawa T, Kanatani Y, Takase B, Kikuchi M, Maehara T. Controlled releases of FGF-2 and paclitaxel from chitosen hydrogels and their subsequent effects on wound repair, angiogenesis and tumor growth. Curr Drug Deliv. 2008;3(4):351–8.
9. Van Tomme SR, Storm G, Hennink WE. In situ gelling hydrogels for pharmaceutical and biomedical applications. Int J Pharm. 2008;355:1–18.
10. Yu L, Ding J. Injectable hydrogels as unique biomedical materials. Chem Soc Rev. 2008;37:1473–81.
11. He C, Kim S, Lee DS. In situ gelling stimuli-sensitive block co-polymer hydrogels for drug delivery. J Control Release. 2008;127:189–207.
12. Han, CK, Bae YH. Inverse thermally-reversible gelation of aqueous N-isopropylacrylamide co-polymer solutions. Polymer. 1998;39:2809–14.
13. Bae Y, Vernon B, Han CK, Kim SW. Extracellular matrix for a rechargeable cell delivery system. J Control Release. 1998;53:249–58.
14. Ruel-Gaiepy E, Leroux JC. In situ-forming hydrogels-review of temperature-sensitive systems. Eur J Pharma Biopharma. 2004:58:409–26.
15. Gappa H, Baudys M, Koh JJ, Kim SW, Bae YH. The effect of zinc-crystallized glucagon-like peptide-1 on insulin secretion of macroencapsulated pancreatic islets. Tissue Eng. 2001:7:35–44.

16. Na K, Park JH, Kim SW, Sun BK, Woo DG, Chung HM, Park KH. Delivery of dexamethasone, ascorbate and growth factor (TGF β-3) in thermo-reversible hydrogel constructs embedded with rabbit chondrocytes. Biomaterials. 2006:27:5951–7.
17. Liu YY, Shao YH, Lu J. Preparation, properties and controlled release behaviours of pH-induced thermosensitive amphiphilic gels. Biomaterials. 2006:27:4016–24.
18. Wadajkar AS, Koppolu B, Rahimi M, Nguyen KT. Cytotoxic evaluation of N-isopropylacrylamide monomers and termperature-sensitive poly(N-isopropylacrylamide) nanoparticles. J Nanoparticle Res. 2009;11(6):1375–82.
19. Paavola A, Kilpelainen I, Yliruusi J, Rosenberg P. Controlled release injectable liposomal gel of ibuprofen for epidural analgesia. Int J Pharm. 2000:199:85–93.
20. Blonder JM, Baird L, Fulfs JC, Rosenthal GJ. Dose-dependent hyperlipidemia in rabbits following administration of poloxamer 407 gel. Life Sci. 1999:65:261–6.
21. Cohn D, Lando G, Sosnik A, Garty S, Levi A. PEO-PPO-PEO based poly(ether ester urethane)s as degradable, reverse thermo-responsive multiblock co-polymers. Biomaterials. 2006:27:1718–27.
22. Moghimi SM, Hunter AC. Poloxamers and polyamines in nanoparticle engineering and experimental medicine. Trends Biotechnol. 2000:18:412–20.
23. Sarkar N. Thermal gelation properties of methyl and hydroxypropyl methylcellulose. J Appl Polym Sci. 1979:24:1073–87.
24. Nystrom B, Walderhaug H, Hansen FK. Rheological behaviour during thermoreversible gelation of aqueous mixtures of ethyl(hydroxyethyl)cellulose and surfactants. Langmuir. 1995:11:750–7.
25. Stabenfeldt SE, Garcia AJ, LaPlaca MC. Thermoreversible laminin-functionalized hydrogel for neural tissue engineering. J Biomed Mater Res. 2006:77:718–25.
26. Baumann MD, Kang CE, Stanwick JC, Wang Y, Kim H, Lapitsky Y, Shoichet M. An injectable drug delivery platform for sustained combination therapy. J Control Release. 2009;138:205–13.
27. Csoka G, Gelencser A, Mako A, Marton S, Zelko R, Klebovich I, Antal I. Potential applications of Metolose® in a thermoresponsive transdermal therapeutic system. Int J Pharm. 2007:338:15–20.
28. Nishinari K,Yamatoya K,Shirakawa M. Xyloglucan. In: PhilIps GO,Williams PA, editors. Handbook of hydrocolloids. Cambridge: Woodhead Publishing; 2000. p 247–66.
29. Miyazaki S, Suisha F, Kawasaki N, Shirakawa M, Yamatoya K, Attwood D. Thermally reversible xyloglucan gels as vehicles for rectal drug delivery. J Control Release. 1998;56:75–83.
30. Suisha F, Kawasaki N, Miyazaki S, Shirakawa M, Yamatoya K, Sasaki M, Attwood D. Xyloglucan gels as sustained release vehicles for the intraperitoneal administration of mitomycin C. Int J Pharm. 1998;172:27–32.
31. Miyazaki S, Suzuki S, Kawasaki N, Endo K, Takahashi A, Attwood D. In situ gelling xyloglucan formulations for sustained release ocular delivery of pilocarpine hydrochloride. Int J Pharm. 2001:229:29–36.
32. Takahashi A, Suzuki S, Kawasaki N, Kubo W, Miyazaki S, Loebenberg R, Bachynsky J, Attwood D. Percutaneous absorption of non-steroidal anti-inflammatory drugs from in situ gelling xyloglucan formulations in rats. Int J Pharm. 2002:246:179–186.
33. Prabaharan M. Review Paper: chitosan derivatives as promising materials for controlled drug delivery. J Biomater Appl. 2008;23:5–36.
34. Chitkara D, Shikanov, A, Kumar N, Domb A. Biodegradable injectable in situ depot-forming drug delivery systems. Macromol Biosci. 2008;6:977–90.
35. Ahmadi R, de Bruijn, J. Biocampatibility and gelation of chitosan-glycerol phosphate hydrogels. J Biomed Mater Res A. 2008;86:824–32.
36. Ji Q, Chen XG, Zhao QS, Liu CS, Cheng XJ, Wang LC. Injectable thermosensitive hydrogel based on chitosan and quaternized chitosan and the biomedical properties. J Mater Sci Mater Med. 2009;20(8):1603–10.
37. Chenite A, Buschmann M, Wang D, Chaput C, Kandani N. Rheological characterization of thermogelling chitosan/glycerol-phosphate solutions. Carbohydr Polym. 2001:46:39–47.
38. Chenite A, Chaput C, Wang D, Combes C, Buschmann MD, Hoemann CD, Leroux JC, Atkinson BL, Binette F, Selmani A. Novel injectable neutral solutions of chitosan form biodegradable gels in situ. Biomaterials. 2000:21:2155–61.
39. Ishihara M, Obara K, Nakamura S, Fujita M, Masuoka K, Kanatani Y, Takase B, Hattori H, Morimoto Y, Ishihara M, Maehara T, Kikuchi M. Chitosan hydrogel as a drug delivery carrier to control angiogenesis. J Artif Organs. 2006:9:8–16.
40. Wu J, Wei W, Wang LY, Su ZG, Ma GH. A thermosensitive hydrogel based on quaternized chitosan and poly(ethylene glycol) for nasal drug delivery system. Biomaterials. 2007:13:2220–32.
41. Suh JK, Matthew HW. Application of chitosan-based polysaccharide biomaterials in cartilage tissue engineering: a review. Biomaterials. 2000:21:2589–98.
42. Jin R, Moreira Teixeira LS, Dijkstra PJ, Karperien, M, van Blitterswijk, CA. Injectable chitosan-based hydrogels for cartilage tissue engineering. Biomaterials. 2008;30:2544–51.
43. Tan H, Chu, CR, Payne KA, Marra KG. Injectable in situ forming biodegradable chitosan-hyaluronic acid based hydrogels for cartilage tissue engineering. Biomaterials. 2009;30:2499–506.
44. Hoemann CD, Sun J, Legare A, McKee MD, Ranger P, Buschmann MD. A thermosensitive polysaccharide gel for cell delivery in cartilage repair. Trans Orthop Res Soc. 2001:26:626.
45. Hoemann CD, Sun J, McKee MD, Chevrier A, Rossomacha E, Rivard GE, Hurtig M, Buschmann MD. Chitosan-glyc-

erol phosphate/blood implants elicit hyaline cartilage repair integrated with porous subchondral bone in microdrilled rabbit defects. Osteoarthritis Cartilage. 2007:15:78–89.
46. Molinaro, G, Leroux JC, Damas J, Adam A. Biocompatibility of thermosensitive chitosan-based hydrogels: an in vivo experimental approach to injectable biomaterials. Biomaterials. 2002:23:2717–22.
47. Friedman CD, Costantino PD, Takagi S, ChowLC. Bonesource hydroxyapatite cement: a novel biomaterial for craniofacial skeletal tissue engineering and reconstruction. J Biomed Mater Res. 1998;43:428–432.
48. Moreau J, Xu H. Mesenchymal stem cell proliferation and differentiation on an injectable calcium phosphate-chitosan composite scaffold. Biomaterials. 2009;30:2675–82.
49. Cho MH, Kim KS, Ahn, HH, Kim MS, Kim SH, Khang G, Bong L, Lee, HB. Chitosan gel as an in situ-forming scaffold for rat bone mesenchymal stem cells in vivo. Tissue Engineering. 2008;14:1099–108.
50. Jeong B, Kim SW, Bae YH. Thermosensitive sol-gel reversible hydrogels. Adv Drug Deliv Rev. 2002:54:37–51.
51. Jeong B, Bae YH, Kim SW. Drug release from biodegradable injectable thermosensitive hydrogel of PEG-PLGA-PEG triblock co-polymers. J Control Release. 2000:63:155–63.
52. Jeong B, Bae YH, Kim SW. In situ gelation of PEG-PLGA-PEG triblock co-polymer aqueous solutions and degradation thereof. J Biomed Mater Res. 2000:50:171–7.
53. Gutowska A, Jeong B, Jasionowski M. Injectable gels for tissue engineering. Anatomical Record. 2001:263:342–9.
54. Yu L, Chang GT, Zhang H, Ding JD. Injectable block copolymer hydrogels for sustained release of a PEGylated drug. Int J Pharma. 2008;348:95–106.
55. Bhattarai N, Ramay HR, Gunn J, Matsen FA, Zhang M. PEG-grafted chitosan as an injectable thermosensitive hydrogel for sustained protein release. J Control Release. 2005:103:609–24.
56. Zentner GM, Rathi R, Shih C, McRea JC, Seo MH, Oh H, Rhee BG, Mestecky J, Moldoveanu Z, Morgan M, Weitman S. Biodegradable block co-polymers for drug delivery of proteins and water-insoluble drugs. J Control Release. 2001:72:203–15.
57. Gong C, Shi S, Dong P, Kan B, Gou M, Wang X, Li X, Luo F, Zhao X, Wei Y, Qian Z. Synthesis and characterization of PEG-PCL-PEG thermosensitive hydrogel. Int J Pharm. 2009;365:89–99.
58. Shim WS, Kim JH, Park H, Kim K, Kwon IC, Lee DS. Biodegradability and biocompatibility of a pH- and thermosensitive hydrogel formed from a sulfonamide-modified poly(ε-caprolactone-co-lactide)–poly(ethylene glycol)–poly(ε-caprolactone-co-lactide) block co-polymer. Biomaterials. 2006:27:5178–85.
59. Chilkoti A, Christensen T, MacKay JA. Stimulus responsive elastin biopolymers: applications in medicine and biotechnology. Curr Opin Chem Biol. 2006:10:652–7.
60. Ge X, Filipe CD. Simultaneous phase transitio2111n of ELP tagged molecules and free ELP: an efficient and reversible capture system. Biomacromolecules. 2006:7:2475–8.
61. Gibson KF, Kernohan WG. Lasers in medicine--a review. J Med Eng Tech. 1995;17:51–7.
62. Kaur M, Srivastava AK. Photopolymerization: a review. J Macromol Sci Poly Rev. 2002;C42:481–512.
63. Nguyen KT, West JL. Photopolymerizable hydrogels for tissue engineering applications. Biomaterials. 2002;23:4307–14.
64. Kumar GS, Neckers DC. Photochemistry of azobenzene-containing polymers. Chem Rev. 1989;89:1915–25.
65. 65. Minkin VI. Photo-, thermo-, solvato-, and electrochromic spiroheterocyclic compounds. Chem Rev. 2004;104:2751–76.
66. Smets G. Photochemical reactions in polymeric systems. Pure Appl Chem. 1975;42:509–26.
67. Deshmukh S, Bromberg L, Smith KA, Hatton TA. Photoresponsive behavior of amphiphilic co-polymers of azobenzene and N,N-dimethylacrylamide in aqueous solutions. Langmuir. 2009;25:3459–66.
68. El Halabieh RH, Mermut O, Barrett CJ. Using light to control physical properties of polymers and surfaces with azobenzene chromophores. Pure Appl Chem. 2004;76:1445–65.
69. Tomer R, Florence AT. Photo-responsive hydyrogels for potential responsive release applications. Int J Pharma. 1993;99:R5–8.
70. Suzuki T, Shinkai S, Sada K. Supramolecular crosslinked linear poly(trimethylene iminium trifluorosulfonimide) polymer gels sensitive to light and thermal stimuli. Adv Mater. 2006;18:1043–6.
71. Pouliquen G, Amiel C, Tribet C. Photoresponsive viscosity and host-guest association in aqueous mixtures of polycyclodextrin with azobenzene-modified poly(acrylic)acid. J Phys Chem B. 2007;111:5587–95.
72. Patnaik S, Sharma AK, Garg BS, Gandhi RP, Gupta KC. Photoregulation of drug release in azo-deztran nanogels. Int J Pharma. 2007;342:184–93.
73. Zhao Y. Rational design of light-controllable polymer micelles. Chem Rec. 2007;7:286–94.
74. Wang G, Tong X, Zhao Y. Preparation of azobenzene-containing amphiphilic diblock co-polymers for light-responsive micellar aggregates. Macromolecules. 2004;37:8911–7.
75. Kuiper JM, Engberts JB. H-aggregation of azobenzene-substituted amphiphiles in vesicular membranes. Langmuir. 2004;20:1152–60.
76. Bisby RH, Mead C, Morgan CG. Active uptake of drugs into photosensitive liposomes and rapid release on UV photolysis. Photochem Photobiol. 2000;72:57–61.
77. Liu X, Jiang M. Optical switching of self-assembly: micellization and micelle-hollow-sphere transition of hydrogen-bonded polymers. Angew Chem Int Ed. 2006;45:3846–50.
78. Sakai H, Matsumura A, Yokoyama S, Saji T, Abe M. Photochemical switching of vesicle formation using

an azobenzene-modified surfactant. J Phys Chem B. 1999;103:10737–40.
79. Kunitake T, Nakashima N, Shimomura M, Okahata Y, Kano K, Ogawa T. Unique properties of chromophore-containing bilayer aggregates: enhanced chirality and photochemically induced morphological change. J Am Chem Soc. 1980;120:6642–4.
80. Zhao H, Sanda F, Masuda T. Stimuli-responsive conjugated polymers. Synthesis and chiroptical properties of polyacetylene carrying L-glutamic acid and azobenzene in the side chain. Polymer. 2006;47:2596–602.
81. Ueno A, Takahashi K, Anzai J, Osa T. Photocontrol of polypeptide helix sense by cis-trans isomerism of side-chain azobenzene moieties. J Am Chem Soc. 1981;103:6410–5.
82. Lustig SR, Everlof GJ, Jaycox GD. Stimuli-responsive polymers. 5. Azobenzene modified polyaramides containing atropisomeric binaphthyl linkages: tuning chiroptical behavior with light and heat. Macromolecules. 2001;34:2364–72.
83. Jaycox GD. Stimuli-responsive polymers. VIII. Polyesters and poly(ester amides) containing azobenzene and chiral binaphthylene segments: highly adaptive materials endowed with light-, heat-, and solvent-regulated optical rotatory power. J Polym Sci Part A: Polym Chem. 2006;44:207–18.
84. Jaycox GD. Azobenzene modified poly(aryl ether ketone amide)s. I. Synthesis and physical properties. Polymer. 1998;39:2589–96.
85. Khoukh S, Oda R, Labrot T, Perrin P, Tribet C. Light-responsive hydrophobic association of azobenzene-modified poly(acrylic acid) with neutral surfactants. Langmuir. 2007;23:94–104.
86. Liu J, He Y, Wang X. Azo polymer colloidal spheres containing different amounts of functional groups and their photoinduced deformation behavior. Langmuir. 2008;24:678–82.
87. Khoukh S, Tribet C, Perrin P. Screening physicochemical parameters to tuning reversible light-triggered control of emulsion type. Colloid and Surfaces A: Physicochem Eng Aspects. 2006;288:121–30.
88. Edahiro J, Sumaru K, Tada Y, Ohi K, Takagi T, Kameda M, Shinbo T, Kanamori T, Yoshimi Y. In situ control of cell adhesion using photoresponsive culture surface. Biomacromolecules. 2005;6:970–4.
89. Sugiura S, Szilágyi A, Sumaru K, Hattori K, Takagi T, Filipcsei G, Zrínyi M, Kanamori T. On-demand microfluidic control by micropatterned light irradiation of a photoresponsive hydrogel sheet. Lab Chip. 2009;9:196–8.
90. Gong C, Lam MHW, Yu H. The fabrication of a photoresponsive molecularly imprinted polymer for the photoregulated uptake and release of caffeine. Adv Funct Mater. 2006;16:1759–67.
91. Gong C, Wong KL, Lam MHW. Photoresponsive molecularly imprinted hydrogels for the photoregulated release and uptake of pharmaceuticals in the aqueous media. Chem Mater. 2008;20:1353–8.
92. Chen J, Serizawa T, Komiyama M. Peptides recognize photoresponsive targets. Angew Chem Int Ed. 2009;48:2917–20.
93. Shimoboji T, Ding ZL, Hoffman AS. Photoswitching of ligand association with a photoresponsive polymer-protein conjugate. Bioconj Chem. 2002;13:915–9.
94. Yu Y, Nakano M, Ikeda T. Photomechanics: directed bending of a polymer film by light. Nature. 2003;425:145.
95. Barrett CJ, Mamiya J, Yager KG, Ikeda T. Photo-mechanical effects in azobenzene-containing soft materials. Soft Matter. 2007;3:1249–61.
96. Yager KG, Barrett CJ. Photomechanical surface patterning in azo-polymer materials. Macromolecules. 2006;39:9320–6.
97. Yuan W, Jiang G, Wang J, Wang G, Song Y, Jiang L. Temperature/light dual-responsive surface with tunable wettability created by modification with an azobenzene-containing co-polymer. Macromolecules. 2006;39:1300–3.
98. Zheng Y, Micic M, Mello SV, Mabrouki M, Andreopoulos FM, Konka V, Pham SM, Leblanc RM. PEG-based hydrogel synthesis via the photodimerization of anthracene groups. Macromolecules. 2002;35:5228–34.
99. Bouas-Laurent H, Castellan A, Desvergne J, Lapouyade R. Photodimerization of anthracenes in fluid solutions: (part. 2) mechanistic aspects of the photocycloaddition and of the photochemical and thermal cleavage. Chem Soc Rev. 2001;30:248–63.
100. Trenor SR, Shultz AR, Love BJ, Long TE. Coumarins in polymers: From light harvesting to photo-cross-linkable tissue scaffolds. Chem Rev. 2004;104:3059–77.
101. Baret V, Gandini A, Rousset E. Photodimerization of heteroarylene-vinylenes. J Photochem Photobio A: Chemistry. 1997;103:169–75.
102. Fang SW, Timpe HJ, Gandini A. Photocrosslinkable polymers bearing pendant conjugated heterocyclic chromophores. Polymer. 2002;43:3505–10.
103. Gandini A, Belgacem MN. Furans in polymer chemistry. Prog Poylm Sci. 1997;22:1203–379.
104. Choi DH, Oh SJ. Photochemical reactions of a dimethyacrylate compound containing a chalcone moiety in the main chain. Eur Poylm J. 2002;38:1559–64.
105. Choi DH, Cha YK. Photo-alignment of low-molecular mass nematic liquid crystals on photoreactive poyimide and polymethacrylate film by irradiation of a linearly polarized UV light. Polym Bull. 2002;48:373–80.
106. Harmon ME, Kuckling D, Frank CW. Photo-cross-linkable PNIPAAm co-polymers. 2. Effects of constraint on temperature and pH-responsive hydrogel layers. Macromolecules. 2003;36:162–72.
107. Decker C, Bianchi C. Photocrosslinking of a maleimide functionalized polymethacrylate. Polym Int. 2003;52(5):722–32.
108. Ishigama T, Murata T, Endo T. The solution photodimerization of (E)-p-nitrocinnamates. Bull Chem Soc Jap. 1976;49:3578–83.
109. Investigation by two-dimensional NMR of the structure and stereochemistry of a methyl p-nitrocinnamate photodimer. J Org Chem. 1988;5:895–6.

110. Greene FD, Misrock SL, Wolfe JRJ. The structure of anthracene photodimers. J Am Chem Soc. 1955;77:3852–5.
111. Bouas-Laurent H, Castellan A, Desvergne J, Lapouyade R. Photodimerization of anthracene in fluid solution: structural aspects. Chem Soc Rev. 2000;29:43–55.
112. Peppas NA, Keys KB, Torres-Lugo M, Lowman AM. Poly(ethylene glycol)-containing hydrogels in drug delivery. J Control Release. 1999 Nov 1;62(1-2):81-7.
113. Andreopoulos FM, Deible CR, Stauffer MT, Weber SG, Wagner WR, Backman EJ, Russell AJ. Photoscissable hydrogel synthesis via rapid photopolymerization of novel PEG-based polymers in the absence of photoinitiator. J Am Chem Soc. 1996;118:6235–40.
114. Andreopoulos FM, Beckman EJ, Russell AJ. Light-induced tailoring of PEG-hydrogel properties. Biomaterials. 1998;19:1343–52.
115. Andreopolous FM, Roberts MJ, Bentley MD, Harris JM, Beckman EJ, Russell AJ. Photoimmobilization of organophosphorus hydrolase within a PEG-based hydrogel. Biotechnol Bioeng. 1999;65:579–88.
116. Andreopoulos FM, Beckman EJ, Russell AJ. Photoswitchable PEG-CA hydrogels and factors that affect their photosensitivity. J Polym Sci A: Polym Chem. 2000;38:1466–76.
117. Zheng Y, Andreopoulos FM, Micic M, Huo Q, Pham SM, Leblanc RM. A novel photoscissile poly(ethylene glycol)-based hydrogel. Adv Funct Mater. 2001;11:37–40.
118. Micic M, Zheng Y, Moy V, Zhang X-, Andreopolous M, Leblanc RM. Comparative studies of surface topography and mechanical properties of a new, photo-switchable PEG-based hydrogel. Colloids and Surfaces B: Biointerfaces. 2002;27:147–58.
119. Andreopoulos FM, Persaud I. Delivery of basic fibroblast growth factor (bFGF) from photoresponsive hydrogel scaffolds. Biomaterials. 2006;27:2468–76.
120. Trenor SR, Long TE, Love BJ. Photoreversible chain extension of poly(ethylene glycol). Macromol Chem Phys. 2004;205:715–23.
121. Gattás-Asfura KM, Weisman E, Andreopoulos FM, Micic M, Muller B, Sirpal S, Pham SM, Leblanc RM. Nitrocinnamate-functionalized gelatin: synthesis and "smart" hydrogel formation via photo-cross-linking. Biomacromolecules. 2005;6:1503–9.
122. Nowakowska M, Zapotoczny S, Sterzel M, Kot E. Novel water-soluble photosensitizers from dextrans. Biomacromolecules. 2004;5:1009–14.
123. Kuang GC, Ji Y, Jia XR, Li Y, Chen EQ, Zhang ZX, Wei Y. Photoresponsive organogels: an amino acid-based dendron functionalized with p-nitrocinnamate. Tetrahedron. 2009;65:3496–501.
124. Chujo Y, Sada K, Nomura R, Naka A, Saegusa T. Photogelation and redox properties of anthracene-disulfide-modified polyoxazolines. Macromolecules. 1993;26:5611–4.
125. Goldbach JT, Russell TP, Penelle J. Synthesis and thin film characterization of poly(styrene-block-methyl methacrylate) containing an anthracene dimer photocleavable junction point. Macromolecules. 2002;35:4271–6.
126. Goldbach JT, Lavery KA, Penelle J, Russell TP. Nano- to macro-sized heterogeneities using cleavable diblock copolymers. Macromolecules. 2004;37:9639–45.
127. Nakayama Y, Matsuda T. Novel surface fixation technology of hydrogel based on photochemical method: Heparin-immobilized hydrogelated surface. J Polym Sci Part A: Polym Chem. 1993;31:977–82.
128. Connal LA, Vestberg R, Hawker CJ, Qiao GG. Fabrication of reversibly crosslinkable, 3-dimensionally conformal polymeric microstructures. Adv Funct Mater. 2008;18:3315–22.
129. Mizutani M, Matsuda T. Photocurable liquid biodegradable co-polymers: In vitro hydrolytic degradation behaviors of photocured films of coumarin-endcapped poly(Î-caprolactone-co-trimethylene carbonate). Biomacromolecules. 2002;2:249–55.
130. Mizutani M, Matsuda T. Liquid photocurable biodegradable co-polymers: In vivo degradation of photocured poly(ε-caprolactone-co-trimethylene carbonate). J Biomed Mater Res. 2002;61(1):53–60.
131. Lendlein A, Jiang H, Junger O, Langer R. Light-induced shape-memory polymers. Nature. 2005;434:879–82.
132. He J, Zhao Y, Zhao Y. Photoinduced bending of a coumarin-containing supramolecular polymer. Soft Matter. 2009;5:308–10.
133. He J, Tong X, Zhao Y. Photoresponsive nanogels based on photocontrollable cross-links. Macromolecules. 2009;42:4845–52.
134. Jiang J, Qi B, Lepage M, Zhao Y. Polymer micelles stabilization on demand through reversible photo-crosslinking. Macromolecules. 2007;40:790–2.
135. Zhao Y, Bertrand J, Tong X, Zhao Y. Photo-cross-linkable polymer micelles in hydrogen-bonding-built layer-by-layer films. Langmuir. 2009;25:13151–7.
136. Yuan X, Fischer K, Schartl W. Photocleavable microcapsules built from photoreactive nanospheres. Langmuir. 2005;21:9374–80.
137. Shi D, Matsusaki M, Akashi M. Unique size-change behavior of photo-crosslinked cinnamic acid derivative nanparticles during hydrolytic degradation. Macromol Biosci. 2009;9:248–55.
138. Sirpal S, Gattas-Asfura KM, Leblanc RM. A photodimerization approach to crosslink and functionalize microgels. Coll Surf B Bioint. 2007;58:116–20.
139. Yuan X, Fischer K, Schartl W. Reversible cluster formations of colloidal nanospheres by interparticle photodimerization. Adv Funct Mater. 2004;14:457–63.
140. Baldursdottir SG, Kjoniksen AL, Karlsen J, Nystrom B, Roots J, Tonnesen HH. Riboflavin-photosensitized changes in aqueous solutions of alginate: rheological studies. Biomacromolecules. 2003;4:429–36.
141. Baldursdottir SG, Kjoniksen AL. Rheological characterization and turbidity of riboflavin-photosensitized changes in alginate/GDL systems. Eur J Pharma Biopharma.

2005;59:501–10.
142. Stern R, Kogan G, Jedrzejas MJ, Soltes L. The many ways to cleave hyaluronan. Biotechnol Adv. 2007;25(6):537–57.
143. Frati E, Khatib A, Front P, Panasyuk A, Aprile F, Mitrovic DR. Degradation of hyaluronic acid by photosensitized riboflavin in vitro. Modulation of the effect by transition metals, radical quenchers, and metal chelators. Free Radic Biol Med. 1997;22(7):1139–44.
144. J. B. Folmer B, Cavini E. Photo-induced depolymerization of reversible supramolecular polymers. Chem Commun. 1998;1998(17):1847–8.
145. Kloxin AM, Kasko AM, Salinas CN, Anseth KS. Photodegradable hydrogels for dynamic tuning of physical and chemical properties. Science. 2009;324:59–63.
146. Jiang J, Tong X, Zhao Y. A new design for light-breakable polymer micelles. J Am Chem Soc. 2005;127:8290–1.
147. Johnson JA, Finn MG, Koberstein JT, Turro NJ. Synthesis of photocleavable linear macromonomers by ATRP and star macromonomers by a tandem ATRP-click reaction: precursors to photodegradable model networks. Macromolecules. 2007;40:3589–98.
148. Sieczkowska B, Millaruelo M, Masserschmidt M, Voit B. New photolabile functional polymers for patterning onto gold obtained by clock chemistry. Macromolecules. 2007;40:2361–70.
149. Brown AA, Azzaroni O, Huck WTS. Photoresponsive polymer brushes for hydrophilic patterning. Langmuir. 2009;25:1744–9.
150. Jiang J, Tong X, Morris D, Zhao Y. Toward photocontrolled release using light-dissociable block co-polymer micelles. Macromolecules. 2006;39:4633–40.
151. Spratt T, Bondurant B, O'Brien DF. Rapid release of liposomal contents upon photoinitiated destabilization with UV exposure. Biochim Biophys Acta. 2003;1611:35–43.
152. Bondurant B, Mueller A, O'Brien DF. Photoinitiated destabilization of sterically stabilized liposomes. Biochim Biophys Acta. 2001;1511:113–22.
153. Zhang ZY, Smith BD. Synthesis and characterization of NVOC-DOPE, a caged photoactivatable derivative of dioleoylphosphatidylethanolamine. Bioconj Chem. 1999;10:1150–2.
154. Burakowska E, Zimmerman SC, Haag R. Photoresponsive crosslinked hyperbranched polyglycerols as smart nanocarriers for guest binding and controlled release. Small. 2009;5:2199–204.
155. McCoy CP, Rooney C, Edwards CR, Jones DS, Gorman SP. Light-triggered molecule-scale drug dosing devices. J Am Chem Soc. 2007;129:9572–3.
156. Levrand B, Herrmann A. Light-induced controlled release of fragrance aldehydes from. 1-alkoxy–9,10-anthraquinones for applications in functional perfumery. Flavour and Fragrance Journal. 2006;21:400–9.
157. Eastoe J, Sanchez-Dominguez M, Cumber H, Burnett G, Wyatt P. Photoresponsive microemulsions. Langmuir. 2003;19:6579–81.
158. Tian F, Yu Y, Wang C, Yang S. Consecutive morphological transitions in nanoaggregates assembled from amphiphilic random co-polymer via water-driven micellization and light-triggered dissociation. Macromolecules. 2008;41:3385–8.
159. Hartner S, Kim HC, Hampp N. Phototriggered release of photolabile drugs via two-photon absorption-induced cleavage of polymer-bound dicoumarin. J Polym Sci Part A: Polym Chem. 2007;45:2443–52.
160. Misu M, Furukawa H, Kwon HJ, Shikinaka K, Kakugo A, Satoh T, Osada Y, Gong JP. Photoinduced in situ formation of various F-actin assemblies with a photoresponsive polycation. J Biomed Mater Res. 2009;89A:424–31.
161. Mamada A, Tanaka T, Kungwachakun D, Irie M. Photo-induced phase transtion of gels. Macromol. 1990;23:1517–9.
162. Irie M, Kunwatchakun D. Photoresponsive polymers. 8. Reversible photostimulated dilation of polyacrylamide gels having triphenylmethan leuco derivatives. Macromolecules. 1986;19:2476–80.
163. Suzuki A, Tanaka T. Phase transition on polymer gels induced by visible light. Nature. 1990;346:345–7.
164. Gorelikov I, Field LM, Kumacheva E. Hybrid microgels photoresponsive in the near-infrared spectral range. J Am Chem Soc. 2004;126:15938–9.
165. Budhlall BM, Marquez M, Velev OD. Microwave, photo- and thermoally responsive PNIPAm-gold nanparticle microgels. Langmuir. 2008;24:11959–66.
166. Akiyama H, Tamaoki N. Synthesis and photoinduced phase transitions of poly(N-isopropylacrylamide) derivative functionalized with terminal azobenzene units. Macromolecules. 2007;40:5129–32.
167. Akiyama H, Tamaoki N. Polymers derived from N-isopropylacrylamide and azobenzene-containing acrylamides: photoresponsive affinity to water. J Polym Sci Part A: Polym Chem. 2004;42:5200–14.
168. Garcia A, Marquez M, Cai T, Rosario R, Hu Z, Gust D, Hayes M, Vail SA, Park CD. Photo-, thermally, and pH-responsive microgels. Langmuir. 2007;23:224–9.
169. Ivanov AE, Eremeev NL, Wahlund P-, Galaev IY, Mattiasson B. Photosensitive co-polymer of N-isopropylacrylamide and methacryloyl derivative of spyrobenzopyran. Polymer. 2002;6;43(13):3819–23.
170. Hu X, Zheng PJ, Zhao XY, Li L, Tam KC, Gan LH. Preparation, characterization and novel photoregulated rheological properties of azobenzene functionalized cellulose derivatives and their alpha-CD complexes. Polymer. 2004;42:6219–25.
171. Zheng PJ, Wang C, Hu X, Tam KC, Li L. Supramolecular complexes of azocellulose and alpha-cyclodextrin: isothermal titration calorimetric and spectroscopic studies. Macromolecules. 2005;38:2859–64.
172. Zheng P, Hu X, Zhao X, Li L, Tam KC, Gan LH. Photoregulated sol-gel transition of novel azobenzene-functionalized hydroxypropyl methylcellulose and its alpha-cyclodextrin complexes. Macromol Rapid Commun. 2004;25:678–82.

173. Luo C, Zuo F, Zheng Z, Ding X, Peng Y. Temperature/light dual-responsive inclusion complexes of alpha-cyclodextrins and azobenzene-containing polymers. J Macromol Sci Part A. 2008;45:364–71.
174. Jiang X, Wang R, Ren Y, Yin J. Responsive polymer nanoparticles formed by poly(ether amine) containing coumarin units and a poly(ethylene oxide) short chain. Langmuir. 2009;25:9629–32.
175. Nagata M, Yamamoto Y. Synthesis and characterization of photocrosslinked poly(e-caprolactone)s showing shape-memory properties. J Polym Sci Part A: Polym Chem. 2009;47:2422–33.

Critical Reviews™ in Biomedical Engineering, 38(6):511–531 (2010)

Ablation of Chronic Total Occlusions Using Kilohertz-Frequency Mechanical Vibrations in Minimally Invasive Angioplasty Procedures

G.B. McGuinness,[1]* M.P. Wylie,[2,3] & G.P.Gavin[2,3]

[1]School of Mechanical and Manufacturing Engineering, Dublin City University, Dublin 9, Ireland; [2]School of Manufacturing and Design Engineering, Dublin Institute of Technology, Dublin 1, Ireland; and [3]Biomedical Devices and Assistive Technologies Research Group, Dublin Institute of Technology, Dublin 1, Ireland

*Address all correspondence to G.B. McGuinness, School of Mechanical and Manufacturing Engineering, Dublin City University, Dublin 9, Ireland; Tel.: +353 1 7005423; Fax: +353 1 7007148; garrett.mcguinness@dcu.ie.

ABSTRACT: Certain minimally invasive cardiology procedures, such as balloon angioplasty and stent implantation, critically require that the site of an arterial blockage be crossed by an intraluminal guidewire. Plaques resulting in near or totally occluded arteries are known as chronic total occlusions, and crossing them with conventional guidewires is a significant challenge. Among the most promising proposed solutions is the delivery of high-power, low-frequency ultrasonic vibrations to the occlusion site via an intraluminal wire waveguide. The vibrating distal tip of the ultrasound wire waveguide is used to transmit energy to the surrounding plaques, tissues, and fluids to ablate or weaken atherosclerotic plaque. Potential mechanisms of interaction with the plaque and adjacent fluids identified in the literature include: (i) direct contact with the waveguide distal tip, (ii) subcavitational acoustic fluid pressure fluctuations, (iii) cavitation, and (iv) acoustic streaming. We summarize developments in this area over more than two decades, describing experimental methods for device performance characterization, preclinical tests, early clinical investigations, and, later, full clinical trials. The article also reviews theoretical foundations and numerical models suitable for device design and analysis. Finally, important issues for future research and for the development of this technology will be considered.

ABBREVIATIONS

CTO, chronic total occlusion; **NiTi,** nickel-titanium

I. INTRODUCTION

Interventions such as percutaneous transluminal coronary angioplasty and stenting are now widely used to restore blood flow through blocked or partially blocked arteries. These procedures involve threading a catheter and balloon assembly along a guidewire to the blockage site via the arterial lumen. A guidewire is a thin wire with distal geometry specifically designed to facilitate navigation through tortuous and branching vessels to establish a viable path to and across the site of a blockage. It acts as a guide rail for catheters to facilitate the positioning of balloon/stent assemblies prior to deployment. Once in position, the balloon is expanded to disrupt and displace the plaque to restore blood flow through the artery. A wire scaffolding structure, known as a stent, may also be expanded by the action of the balloon, and its purpose is to hold the artery open, preventing recoil once the balloon is removed.

A major prerequisite for the use of these interventional procedures is the ability to first traverse the lesion with an intraluminal guidewire. Depending on the nature of the blockage, this may be extremely difficult or impossible with conventional guidewires. Such cases are often characterized as chronic total occlusions (CTOs), referring to total closure of the vessel. Several novel approaches to overcoming this problem have been proposed. Among the most promising is a technology involving the delivery of high-power, kilohertz-frequency ultrasonic vibrations to the total occlusion site via an intraluminal wire waveguide to assist the progression of the guidewire across the lesion by a process of plaque ablation or disintegration.

1040-8401/10/$35.00

II. ATHEROSCLEROTIC PLAQUES

A. Atherogenesis

Atherosclerosis is a thickening of the arterial medial layer that affects the artery by altering conduit function and reducing blood flow downstream. There are numerous hypotheses describing the origin and progression of the disease, which are beyond the scope of this review. It is known to be a time-dependent process with varying stages of severity,[1-3] as confirmed by extensive autopsy studies.[4] The progression of atherosclerosis can be broadly classified into three stages: the fatty streak, the fibroatheromatous plaque, and the complicated lesion.[5,6]

Of particular interest are the complicated calcified plaques that have a plaque core and cap that are primarily composed of calcified minerals and tough collagen fibers.[5] They are closely associated with CTOs, and the success rate with standard procedures involving the mechanical loading of the lesion tend to be lower because the rigidity and stiffness of the calcified material and tough collagen fibers can resist these applied loads.[5]

B. Mechanical Characteristics of Atherosclerotic Plaques

Plaque mechanical properties are largely responsible for problems experienced with the use of standard interventional procedures for CTOs. As biological materials, these plaques exhibit complex properties in response to mechanical compression or indentation.

Loree et al. excised plaque specimens in the circumferential direction and classified them histologically as cellular (12 samples), hypocellular (9 samples), or calcified (5 samples).[7] The hypocellular plaques were found, on average, to be 1 to 2 times stiffer than cellular plaques, and calcified plaques were 4 to 5 times stiffer than cellular plaques.

Topoleski and Salunke[8] also reported the mechanical behavior of various plaque types as well as contiguous arterial tissue. Atherosclerotic lesions (plaque cap and core) were removed from excised human aortoiliac arteries, classified as calcified (hard), fibrous (medium), or atheromatous (soft) core. The plaque compressive mechanical response was non-linear and could be categorized into three distinct mechanical behavior patterns. The hard calcified and fibrous plaques (containing some calcium) showed a distinct behavior, with a considerably stiffer response under compression, than the medium and soft plaques and the healthy arterial tissue.

C. Complications Associated with Percutaneous Coronary Interventions

CTOs are challenging because they do not lend themselves to access by the guidewire. Guidewire access and ability to cross the lesion is critical to the success of standard dilation procedures and has been reported to be the key success indicator in 80% of cases.[9] In many cases, traditional guidewires are not stiff enough to transverse the tough fibrous cap, and the higher forces applied could carry an associated risk of arterial perforation. Therefore, bypass graft surgery is often required, which is a more invasive and less desirable treatment option. CTOs are a significant cause for coronary bypass graft surgery referral.[10] Not only are there problems traversing the CTO, but there are other issues with safe re-entry into the target lumen. CTOs are known to have microvessels that are formed within the organizing thrombi (a result of endothelial cells invading the fibrin lattice), which may hinder successful dilation by angioplasty.[11]

More advanced lesions can be unstable and potentially hazardous. Death and myocardial infarction may occur during CTO angioplasty by shearing off the collateral circulation, thrombus formation, perforation, arrhythmia, or damaging the proximal epicardial coronary artery or proximal side branches.[12]

Chronic total occlusion of a coronary artery is usually defined as "an obstruction of a native coronary artery for greater than 30 days with no luminal continuity and with thrombolysis in myocardial infarction (TIMI) flow grade 0 or 1."[13] CTOs are generally found in small arteries such as those in the coronary artery tree. Chronic total occlusion of the right coronary artery is identified in approximately 20% of angioplasty patients, and angioplasty of the CTO is attempted in 10% to 15% of all cases.[14] However, standard techniques using contemporary guidewires are unsuccessful in approximately 20% of these cases.[12,15] Success rates in CTOs have

steadily increased over the last 15 years because of greater operator experience, improvements in equipment, and procedural techniques; however, minimally invasive procedures on these types of lesions are the most likely to fail. The majority of failures are due to the inability to successfully pass a guidewire across the lesion (89%).[15] The ability to cross the lesion with a guidewire is the main determinant of interventional success, and if this cannot be achieved, then dilation devices such as balloon angioplasty and stents cannot gain access to the lesion site to reopen the blockage.

While plaques can develop around the entire lumen structure, resulting in a concentric lesion, progression is often eccentric.[16-18] Calcified eccentric lesions pose further complications because during dilation procedures the rigid, calcified side of the artery can potentially remain non-dilated, while the healthy and often thinner side of the arterial wall over-dilates, meaning that mechanical dilation-based interventions with eccentric lesions can be less effective.[16] Thus, standard interventional procedures work best with concentric lesions because the pressure is divided relatively evenly over the lesion. However, dilation of eccentric lesions can be problematic because less-diseased walls become overstretched, increasing the risk of cell necrosis, tearing, and restenosis.

Interventional procedures for these applications should ideally incorporate a means of establishing access to the site by navigating the vascular structure, be capable of disrupting total blockages due to plaque or thrombus so that a guidewire can be advanced, and also be able to selectively damage calcified material while leaving healthy tissue largely unaffected.

D. Recent Clinical Progress

A conventional approach to the problem involves the design of specialized passive guidewires. The attributes of typical commercial devices are summarized in Table 1, and are more comprehensively reviewed elsewhere.[19] Godino et al. list the critical characteristics of a guidewire as "tip load, tip stiffness guidewire flexibility, ability to shape, shaping memory, shaft support, torque transmission, trackability and resistance to tracking of the wire within the occlusion."[19] The majority of current guidewires have a main shaft diameter of 0.35 mm or less. The tips vary from 0.2 to 0.35 mm in diameter, depending on the guidewire. Guidewires fall into two classifications: hydrophilic, which give better tactile feedback, and non-hydrophilic, which have lower resistance and better maneuverability.[19] Guidewires are also classified as soft, intermediate, or stiff. Soft wires are normally used for advancement of the catheter and for

TABLE 1. Variety of Guidewires Showing Shaft and Tip Diameters

Manufacturer	Wire	Shaft Diameter (mm)
Guidant	Whisper	0.35
	Pilot 50	0.35
	Pilot 150 & 200	0.35
	HT Intermediate	0.35
	HT Standard	0.35
	Cross-IT 100-400	0.35
Boston Scientific	Choice PT & PT2	0.35
	PT Graphix & P2	0.35
Cordis	Shinobi & Shinobi Plus	0.35
Medtronic Vascular	Persuader	0.35
	Persuader 9	0.35 (tip diameter 0.27)
Abbott Vascular Asahi	Confianza	0.35
	Confianza Pro (Conquest)	0.35 (tip diameter 0.22)
	Medium	0.35
	Miraclebros	0.35

crossing occlusions with small lumens, intermediate wires are used for recently occluded lesions or tortuous vessels, and stiff wires are used for advancing through CTOs. The suggested method for crossing CTOs is to use a soft or intermediate wire to explore the area. A stiffer guidewire can be used to cross the proximal cap if needed, and an even stiffer guidewire can then be used, again, if necessary, to cross the CTO and penetrate the distal cap. After the occlusion has been crossed, the stiff guidewire is then replaced with a soft guidewire.[19]

Lefevre[20] reports that, during procedures involving CTOs, crossing lesions may take up to 35 minutes and, with the possibility of guidewire changeovers, entire procedures may take 1.5 hours. This has implications for the cost of procedures. A soft guidewire with the capability of crossing CTOs would offer major advantages, reducing the need for changing wires and therefore reducing the time required for crossing CTOs, and thus increasing the likelihood of success for the procedure.

As an alternative approach, therapeutic ultrasound transmitted via wire waveguides has long been identified as a technology with the potential to disrupt atherosclerotic lesions with mechanical vibrations transmitted via long, low-profile, flexible superelastic wire wave guides.[25-25] Several prototype devices have been developed in an attempt to meet this challenge, including 20-kHz ultrasound wire waveguide clinical devices, which deliver high-power, low-frequency intravascular acoustic energy to the sites of CTOs. Early clinical studies on femoral arteries were conducted by Rosenschein et al., and showed significantly larger areas of recanalization than the unactivated control guidewires (mean of 5.9 versus 1.7 mm2; $P < 0.05$).[23] Siegel et al. conducted further clinical tests on human peripheral arteries, also achieving significant reductions in stenosis.[24] Subsequently, Siegel reported initial clinical experience with the use of this technology for coronary artery angioplasty on 19 subjects, with arterial stenosis reduced from 80% ± 12% to 60% ± 18% ($P < 0.001$) with the ultrasound treatment and then to 26% ± 11% ($P < 0.001$) with a follow-up balloon angioplasty.[25]

In early 2005, Flowcardia™ Inc. (Sunnyvale, CA) received approval to market an ultrasonic waveguide (the Crosser© System) in the European Union for the treatment of CTOs following failure to cross by conventional guidewire techniques,[26] and this device was granted US Food and Drug Administration approval in 2007. An angiogram depicting the use of the Flowcardia device is shown in Figure 1, which includes the occluded artery (a), the ball-tipped probe in position (b), the device crossing the lesion (c), and the final lumen (d).

The Flowcardia device, based on a monorail concept, is approved specifically for use on CTOs. The system used is run at a frequency of approximately 20 kHz and with a working length of 146 cm. The ultrasound is transmitted via a nickel-titanium wire waveguide to a 1-mm[en]diameter stainless steel tip. Apart from the tip, the ultrasound wire and guidewire are all housed in a 6 F catheter, which acts as an irrigation system. The system works by moving a guidewire to the point of the lesion, and then advancing the stainless steel tip over the guidewire, fragmenting the lesion as it is pushed forward.

In one trial, 55 CTOs in 53 patients were treated using the system. The device had a success rate of 76%, with no major cardiac events or coronary perforations.[27] In a second study with 30 lesions in 28 patients, success was achieved in 63% of the procedures[26]; however, there was one guidewire perforation with no serious adverse effects and one peri-procedural myocardial infarction.

III. FUNDAMENTALS OF HIGH-POWER, LOW-FREQUENCY THERAPEUTIC ULTRASOUND ANGIOPLASTY

A. Background

In the present context, therapeutic ultrasound refers to the use of high-amplitude, low-frequency ultrasound in clinical interventions.[28] The mechanical effect that this form of ultrasound has on biological tissues was first noted by Conte and de Lorenzi,[29] and its effectiveness is based on the fact that at the right combination of frequency and amplitude, inelastic rigid tissue is vigorously disrupted while elastic tissue can absorb the energy.[24,30]

It was therefore conceived that this form of energy may be useful in the treatment of cardiovas-

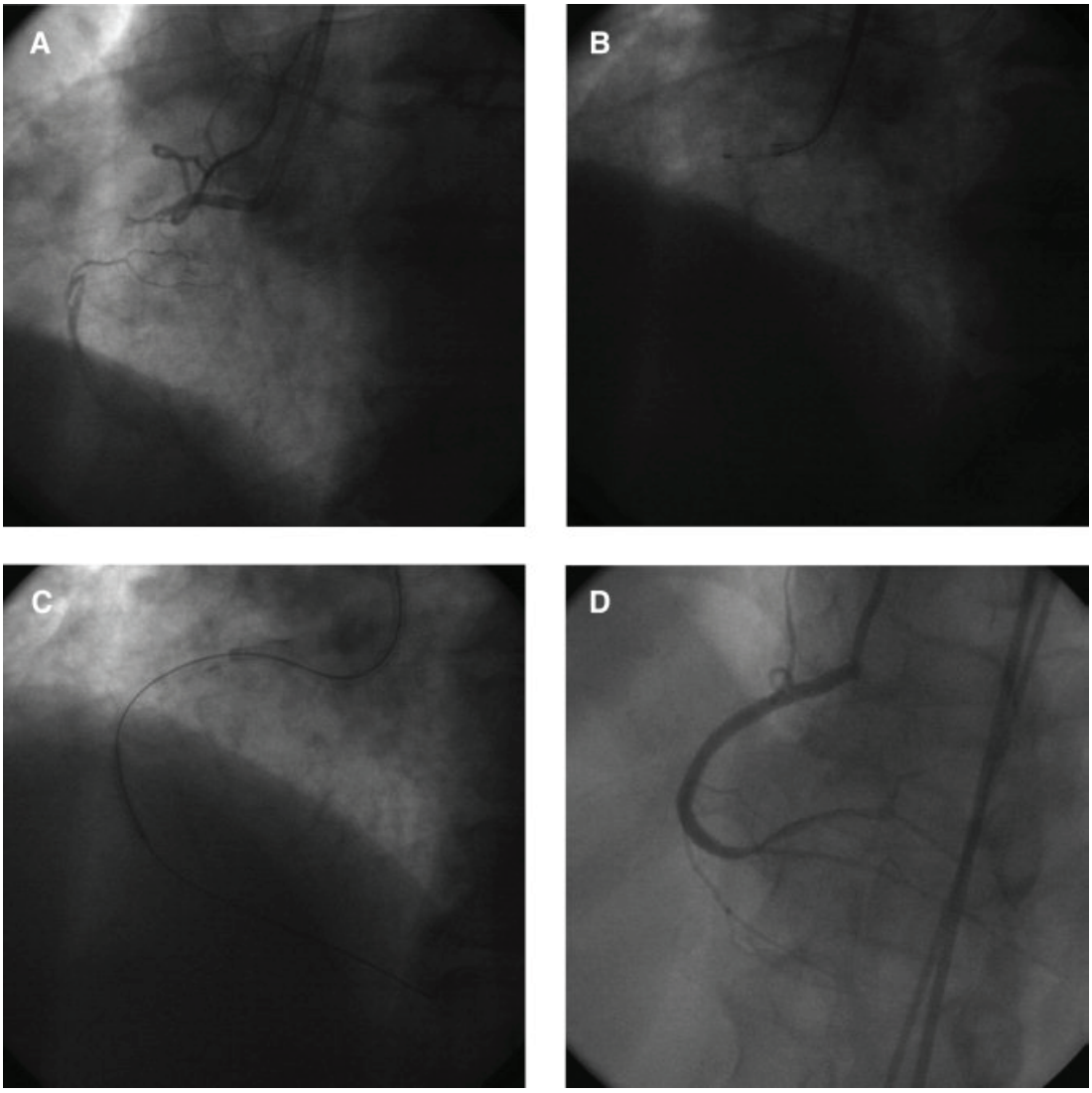


FIGURE 1. Clinical use of the Flowcardia crosser. (a) Preliminary angiogram, (b) introduction of device, (c) crossing of the occlusion, and (d) angiogram of reopened vessel. (From Melzi et al.[26]; used with kind permission of John Wiley & Sons.)

cular disease and could potentially have advantages over standard dilation procedures in the targeting of specific lesions, especially the complicated rigid calcified and fibrous plaques.[28]

Development and testing of experimental devices began as early as the 1970s, but these devices were extremely limited for practical use. Sobbe et al.31 showed that ultrasound delivered through a large-diameter wire probe resulted in a longitudinally vibrating distal tip that had the effect of disrupting blood clots in animals.

During the mid-1980s, particular design issues were addressed by two groups, headed by Siegel and by Rosenschein, with the goal of making an actual working prototype for initial clinical testing and, potentially, for use in trials. No specific design issues appear in the literature, and the majority of this work was focused on the end clinical results. Some general information is reported and is included in this review.[21-23]

Both teams based their design efforts on the system developed by Sobbe et al.,[31] delivering the

ultrasonic waves to the lesion via a wire waveguide. This sets up longitudinal stress waves and a longitudinal peak-to-peak displacement at the distal tip of the wire waveguide with the potential to disrupt both lesions and clots.

B. Ultrasound Generation

To achieve the micron amplitudes and kilohertz frequencies required to cause disruption to lesions, a source capable of generating and transmitting these ultrasonic vibrations was required. Both Siegel et al.[21] and Rosenschein et al.[22,23] describe the use of a piezoelectric transducer as a source for the ultrasound.

The piezoelectric effect is a property of certain classes of crystalline materials, including the natural crystals of quartz, rochelle salt, and tourmaline plus manufactured ceramics such as barium titanate and lead zirconate titanates. When a mechanical pressure is applied to these materials, a voltage is produced that is proportional to the pressure applied. Conversely, when a voltage is applied, the structure changes shape, acting as an electromechanical transducer. These shape changes are usually very small, usually in the order of a few microns.

Dynamic voltages can also be applied, and result in a dynamic displacement or shape change in the material. In this arrangement, the material behaves very much like a mechanical system with resonant frequency characteristics. The transducer's output will be at a maximum at the first resonant frequency, a characteristic of the material and geometric configuration. Using a stack arrangement, as shown in Figure 2a, mechanical amplitudes with a vibration of 0 to 5 μm peak to peak at frequencies less than 100 kHz can be achieved. An ultrasonic generator provides the electrical source to drive the transducer at the resonant frequency of the piezoelectric stack.[32,33]

These types of transducers are used mostly for sonochemistry applications in which agitation of chemical and biological samples is required. The frequencies and amplitudes discussed are chosen for their ability to cause cavitation, a desirable effect in processing chemical solutions. Therefore, transducers used for sonochemistry applications may be adapted for the generation of intravascular therapeutic ultrasound.[21] For sonochemistry and the generation of cavitation, the displacements (at the frequencies used) from the converter (transducer) are still too small for sample processing. An acoustic horn or waveguide is usually attached to the transducer to amplify the displacements.

Acoustic horns are solid metal rods that are designed to couple to the front-end mass of the converter. They achieve an increase in output displacement by two means. First, their geometry is such that the input wave is compressed through a progressively smaller cross-sectional area as it travels the length of the rod, resulting in a larger displacement at the output. This can be clearly seen in both linear and exponentially tapered horns.[33] Second, horns can be manufactured to resonate at the frequency of the ultrasonic converter. Stepped horns, as shown in Figure 2b, appear to be the most useful for adaptation to the natural frequency of the driving transducer and are often the easiest to manufacture.[34]

In most practical applications, a combination of stepped and tapered sections is used in horn design. Horns are manufactured from materials that have high dynamic fatigue strength and low acoustic loss, such as titanium alloys. With a horn attached to the converter output, amplitudes of vibration greater than 150 μm can be achieved at frequencies less than 100 kHz.

Acoustic horns, being of solid metal configuration, lack the flexibility necessary to navigate the tortuous vascular geometry. The idea of using wire waveguides was developed to deliver these ultrasonic peak-to-peak displacements over sufficient lengths and with the flexibility necessary for minimally invasive vascular surgery.[21,22,31]

C. Minimally Invasive Delivery of Ultrasound by Wire Waveguide

Most of the initial work in the area was based on finding methods to deliver this form of ultrasound over the waveguide lengths and small diameters capable of being used in surgical applications. While little detailed description of the exact design and construction methods was given, some general conclusions can be drawn from early design and testing.

Rosenschein et al.[23] described the use of a solid,

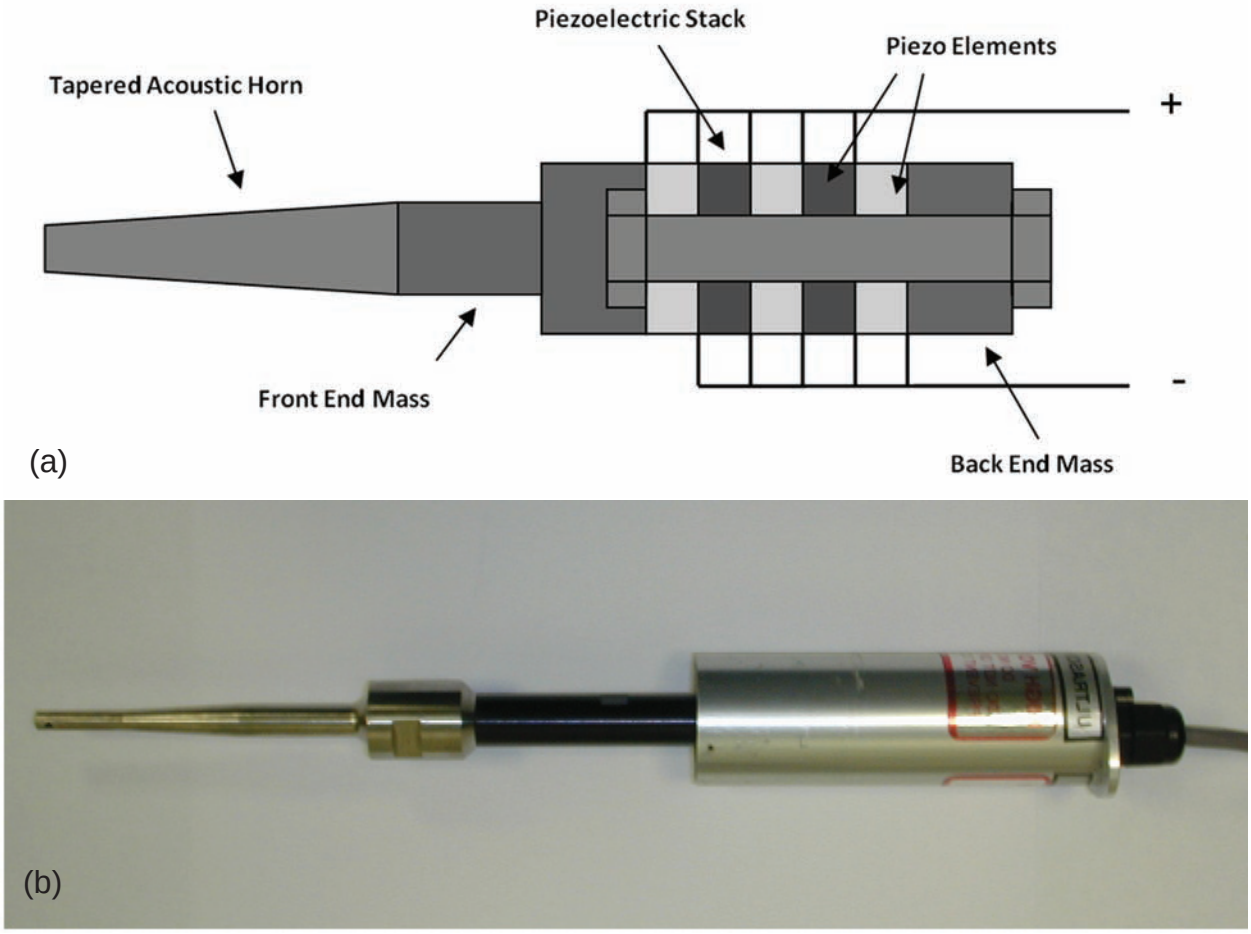


FIGURE 2. (a) Generation of mechanical ultrasonic displacements: converter and acoustic horn. (b) Assembled acoustic horn and converter.

flexible aluminum transmission wire mechanically coupled to the acoustic horn. Details surrounding the coupling method were not included. Fischell et al. describe the use of a solid 1.5 F (≈ 0.5-mm) diameter titanium wire waveguide.[35] This system also appears to have been used by Ariani et al.[36] and Demer et al.[30] These studies, in general, detail the use of solid wires manufactured from aluminum, titanium, or alloys of both to form the transmission member or wire waveguide to transmit the ultrasound from the acoustic horn to the lesion location. This ultrasonic transmission results in longitudinal displacements at the distal tip of the wire waveguide.

In addition, both teams located a ball-tip or enlargement at the distal end of the wire. This increases the surface area in contact with the lesion and the surrounding fluid. Demer et al.[30] described the use of a 2.0-mm[en]diameter ball-tip, and Siegel et al.[25] a 1.7-mm[en]diameter ball-tip in conjunction with the wire waveguide.

D. Mechanical Effect of Wire Waveguide Tip Displacement

Atar et al.[29] suggested that the ultrasonic longitudinal vibration of the wire waveguide distal tip results in four major potentially disruptive events—direct contact ablation, pressure wave components, cavitation, and acoustic micro-streaming—and all of these appear to be related to the distal-tip displacement

amplitude, frequency, and geometry.[37-40] A diagram of the distal section of the wire waveguide and location of disruption mechanisms is shown in Figure 3. Distal tip displacements described in the literature are given in Table 2.

1. Direct Contact Ablation

The oscillating distal tip of the wire waveguide acts as an ultrasonic reciprocating micro-drill. Research in ultrasonic micromachining has shown that it is associated with low material removal rates by microchipping, and is ideal for ablating brittle ceramic materials.[41] Direct contact between the oscillating tip and the plaque in the arterial lumen results in the fragmentation and ablation of the plaque into microscopic particles, and is considered to be one of the major effects of ultrasound angioplasty.[24] This issue was further examined in a study by Siegel et al. on 11 atherosclerotically occluded human coronary arteries that had been excised postmortem.[25] Following the use of a 19.5-kHz ultrasound device with amplitudes of 15 to 30 µm, a resistive pulse particle counter was used to measure the particulate content in the effluent. Particles 2.5 to 80 µm in size were detected, but 99% of particulate content was found to be less than

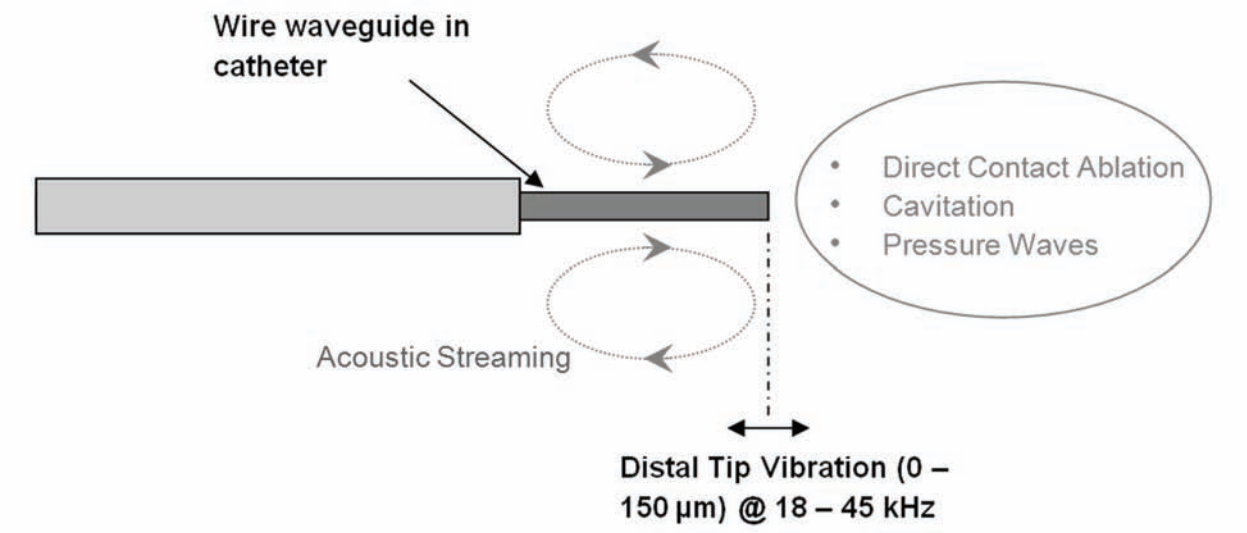


FIGURE 3. Schematic of ultrasonic wire waveguide in catheter and the regions surrounding the longitudinal vibrating distal tip where disruptive mechanisms can occur.

TABLE 2. Waveguide Tip Displacements from the Literature

Study	Frequency of Operation (kHz)	Distal Peak-to-Peak Displacement (µm)	Wire Data
Rosenschein et al.[23]	20	150 ± 25	Aluminum alloy wire 1.6 mm No ball tip
Ariani et al.[36]	20	63.5–11	Titanium wire 0.72 mm 2-mm ball tip
Demer et al.[30]	20	50 ± 25	Titanium wire 0.5 mm 2-mm Ball-Tip
Makin and Everbach[44]	22.5	200/130	Titanium wire 1.98/2.46-mm ball tip

10 µm in diameter. Ernst et al. describe a study using a 21-kHz device with a 52 ± 19-µm amplitude of displacement, and found that 95% of the debris particles were of a diameter under 9 µm (range, 5–12 µm).[42] It also appears that, due to the varying material properties of the plaques, this form of fragmentation technique ablates less distensible rigid calcified and fibrous plaques faster than flexible material such as the healthy arterial wall tissue.[22]

2. Acoustic Pressure Waves and Cavitation

As a result of the direct contact between the oscillating distal tip and the surrounding fluid, an oscillating acoustic pressure field superimposed on the ambient pressure is established around the distal tip. Of particular interest is the fact that if the pressure amplitude is sufficiently high, cavitation in the fluid may occur. Cavitation occurs when on the negative side of a pressure cycle (e.g., when the wire waveguide tip is retracting) suspended gas bubbles in the fluid, in channels within the tissue, or trapped at solid interfaces expand and collapse with the generation of shock waves of sufficient amplitude and frequency. If the acoustic field produces sufficiently low pressures, cavitation may occur. This is considered to be the most powerful destructive mechanism of the ultrasonic ablation.[43]

Burdic[37] has suggested a simplified relationship between pressure and cavitation threshold, discussed in the section called "Distal Tip Interactions" below. Cavitation is a potentially significant erosion event, undesirable in most acoustic applications but necessary in cavitation cleaning baths and sonochemistry.[37,38]

Yock and Fitzgerald[28] concluded that cavitation is a major contributing factor in the disruption of plaque and thrombus, that ultrasound ablation of lesions is only present above the cavitation threshold, and that the rate of disintegration is correlated with the amount of power delivered above this threshold.

Makin and Everbach[44] investigated the acoustic pressures developed by an ultrasonically vibrating wire waveguide submerged in a liquid with acoustic properties similar to blood. The experiment consisted of a 2.46-mm[en]diameter spherical tipped wire oscillating in a cylinder of fluid (peripheral arterial phantom) at 22.5 kHz. Their experiment consisted of an ultrasonic emitter-receiver submerged in a tube within a 355-mm[en]tall acrylic tank with an internal diameter of 203 mm. The tank was filled with water or a glycerine-water mix. The experiment was described as having a 355-mm[en]high tank filled to 345 mm with fluid. Pressures were measured in the range of 12 to 250 mm from the vibrating tip using an acoustic hydrophone; measurements in the vicinity of the tip were restricted due to the limitations of their measuring equipment.

Their results demonstrate two important effects of the ultrasound in vivo:

1. While the investigators were not able to directly measure pressures at the tip, "cavitation activity" was determined using a 20-MHz focused transducer. They concluded that cavitation was evident and related to the distal-tip displacement and the distal tip geometry.

2. A standing wave in the acoustic domain of significant amplitude was also detected. They concluded that this standing wave was caused by an impedance mismatch between the liquid and the surrounding air. This is similar to the tissue-air interface encountered in peripheral limbs in vivo, the result of reflection at the interface of layers/materials of different acoustic impedance. The sound power reflection coefficient is given as[45]:

$$\alpha_r = \left(\frac{\rho_2 c_2 - \rho_1 c_1}{\rho_2 c_2 + \rho_1 c_1} \right)^2 \quad (1)$$

where α_r is the sound power reflection coefficient and c is the speed of sound in fluid medium.

Assuming no energy loss, the reflection of the incident energy must correspond to a unity value of reflection coefficient. The speed of sound for air and blood/tissue was taken as 340 m/s and 1580 m/s, respectively, and with densities of 1.2 kg/m^3 and 1050 kg/m^3, respectively. Using Equation 1 and the material properties shown, the sound/power reflection can be calculated to be 99.9% at the tissue-air interface.

3. Acoustic Streaming

Two general forms of fluid motion are set up around an oscillating sphere in a fluid. The first can be considered as the oscillatory fluid motion very close to the wall of the tip, and the second a unidirectional fluid motion in an external acoustic streaming layer, as shown in Figure 4.[38,39] It is the combined presence of direct contact, acoustic pressure waves, cavitation, and acoustic streaming that led investigators to believe that ultrasound delivered via a wire waveguide could disrupt arterial lesions. Some analytical theory further describing these disruptive mechanisms is discussed later.

E. Selected Examples of the Therapeutic use of Megahertz-Frequency Ultrasound Catheters

There is an extensive body of literature concerning the use of megahertz-frequency ultrasound transducers at a catheter tip to produce a therapeutic benefit, which are considered beyond the scope of this review. It should be noted that these systems employ a megahertz-frequency ultrasound transducer attached to the distal tip of a catheter. They are therefore quite technically different from the kilohertz systems under discussion, which involve external ultrasound generation and transmission along a metal wire waveguide. For further information on megahertz-frequency ultrasound therapy, the reader is directed to recent contributions and reviews in this related field.

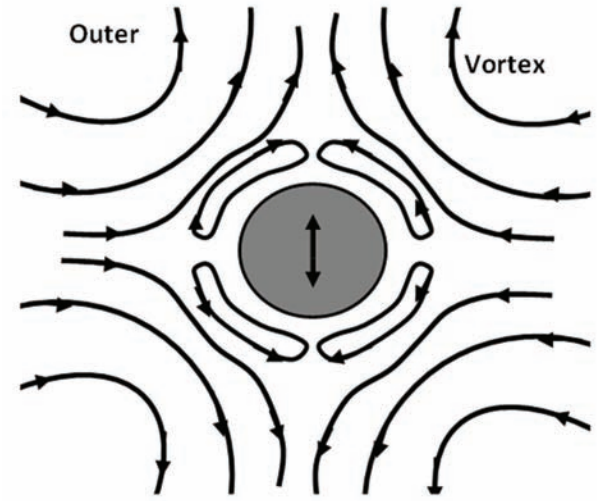


FIGURE 4. General features of acoustic microstreaming near a small vibrating sphere. (Adapted with permission from Nyborg.[38])

For neurosurgery applications, Herickhoff et al. reported the use of megahertz-frequency ultrasound for dual-mode imaging and treatment of brain tumors,[46] and the use of a thermal model to predict temperature rise in surrounding brain tissue.[47]

Lafon et al. developed flexible tubular interstitial applicators for focused ultrasound tissue destruction, whereby the transducers can be introduced into the body and to the proximity of the lesion site in a fashion similar to the catheter systems mentioned previously.[48] They have also used numerical models for optimization of the shape of ultrasound transducers through consideration of thermal lesion predictions and involving solution of the bioheat equation. A review of interstitial devices for minimally invasive thermal ablation by high-intensity ultrasound is also available.[49]

The use of a megahertz-frequency ultrasound catheter for accelerated thrombolysis of deep vein thrombosis in combination with thrombolytic drugs has been described by Parikh et al., with some promising results.[50]

Smikhal et al. previously studied the combination of megahertz-frequency ultrasound in combination with alteplase, a lytic protein for deep vein thrombosis thrombolysis (using an Omnisonics ultrasound systems, which transmits lateral kilohertz vibrations via a metal wire waveguide, as opposed to the axial vibrations that are the main theme of this article).[51]

IV. THEORETICAL FOUNDATIONS

This section focuses on the theoretical mechanics that govern the response of a uniform rod (waveguide) to a harmonic input displacement, as well as the fundamental principles of acoustic pressure field generation by an oscillating sphere (such as a distal ball tip) in a fluid.

A. Wire Waveguide Mechanics

1. Steady-State Vibration of a Uniform Rod

The steady-state analytical solution of motion for an undamped longitudinally vibrating rod, as shown in Figure 5, subjected to a sinusoidal input vibration motion of u(t) =b sin (ω(t)), is shown in Equation 2. This is similar to the condition setup in ultrasound transmitted via a wire waveguide, where the distal tip of the acoustic horn applies a sinusoidal input of particular amplitude and frequency to the proximal end of the wire waveguide. The theoretical developments leading to this solution are available.[52]

$$u(x,t) = b(\cos\frac{\omega x}{c} + \tan\frac{\omega l}{c}\sin\frac{\omega x}{c})\sin\omega t \qquad (2)$$

where *x* is the axial coordinate, denoting position along the rod; *u(x,t)* is the axial displacement of the rod as a function of *x* and time *t*; ω is the frequency of vibration; *l* is the length of the rod; *b* is the amplitude of the applied sinusoidal vibration at x = 0 (as indicated above); and c is the speed of sound in the medium.

The steady-state amplitude of vibration at any point ($0 \leq x \leq l$) can be determined, although this is an undamped solution and therefore limited. The solution can also be used in the determination of resonant response.

The solution for the resonant response from Equation 1 is shown in Equation 3, where f_n is the resonant frequency of vibration for a thin rod of length *l* and *c* is the speed of sound (longitudinal) in the rod material.

$$f_n = \frac{nc}{4l} \qquad n = 1, 3, 5... \qquad (3)$$

Similarly, for a constant frequency, Equation 4 gives us the lengths where resonance occurs (n = 1, 3, 5, 7...) and the lengths where non-resonance occurs (n = 0, 2, 4, 6, 8…). This is more important to the ultrasound wire waveguide because the displacement input from the distal tip of the acoustic horn is at a constant prescribed frequency determined by the ultrasonic generator.

$$l_n = \frac{nc}{4f} \qquad (4)$$

The reader is referred to Steidel[52] for a complete treatment of the principles on which these equations are based.

B. Distal-Tip Interactions

The literature describes ultrasound transmission resulting in vibration of the distal tip of the wire waveguide and the presence of a spherical or near spherical ball-tip enhancing transmission to the fluid.[44]

1. Acoustic Pressure Field Around an Oscillating Sphere

An analytical solution for the pressure field developed in the fluid near an oscillating sphere was proposed by Morse[44] and is shown in Equation 5. Figure 6 illustrates the parameters appearing in this solution. In this solution, the maximum pressure at any point, P_{max}, developed in the field surrounding an oscillating sphere can be determined.

$$P_{\max}(r,\theta) = 2\pi^2\rho R f^2 d_0 \times \frac{R^2|\cos\theta|}{r^2} \qquad (5)$$

Where *R* is the radius of the ball, *f* is the frequen-

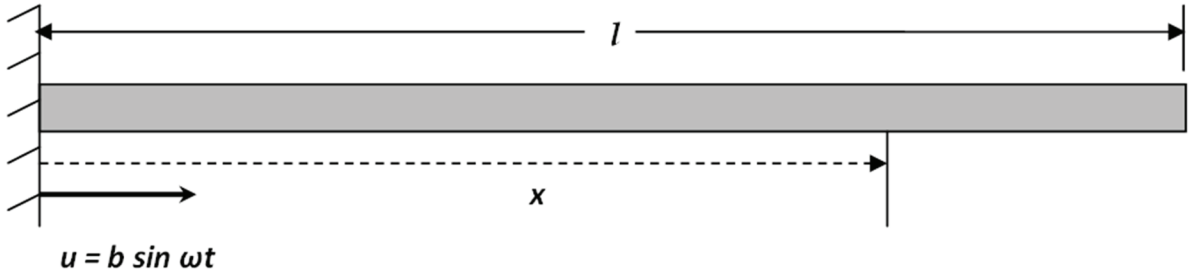


FIGURE 5. Diagram of a uniform rod, of length l, with an input displacement motion of u = b sin ωt. (From Steidel[52]; reproduced with kind permission of the ASME.)

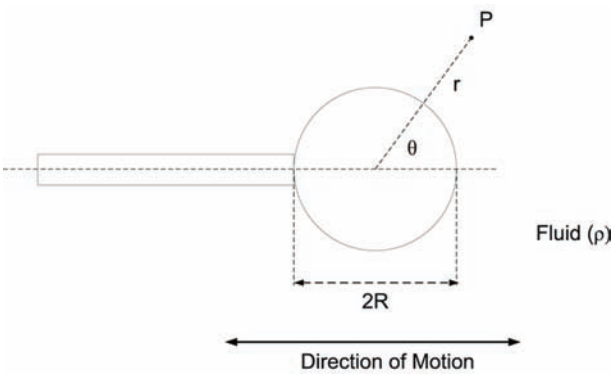


FIGURE 6. Diagram relating to the pressure field developed around an oscillating sphere. (Adapted from Nyborg[38]; originally published by Gavin et al.[56] and reproduced with kind permission of the ASME.)

cy of oscillation, d_0 is the displacement amplitude of oscillation, and ρ is the fluid density. This solution shows that pressure amplitudes are at a maximum on the surface of the sphere and axially ahead of it.

2. Cavitation

From the pressure amplitude field, it is possible to infer when and where cavitation is likely to occur. While the occurrence of cavitation is dependent on numerous factors, such as the temperature and the amount of dissolved gas in the fluid, Burdic[37] suggests the use of a cavitation threshold for the sound intensity, which may be calculated according to Equation 6:

$$T = \frac{P_{\max}^2}{2\rho c} \quad (6)$$

Where T is the sound intensity, P_{max} is the maximum pressure, ρ is the fluid density, and c is the speed of sound in the media.

According to Perkins,[53] this threshold intensity for degassed water at room temperature is approximately between 2 and 3 W/cm^2 in the frequency range of 20 to 30 kHz. It is worth noting here that the acoustic intensity required to cause cavitation increases significantly above a frequency of 100 kHz, and is one of the main reasons that sonochemistry and therapeutic ultrasound is performed at low-frequency ultrasound bands of between 20 and 45 kHz.

3. Acoustic Streaming

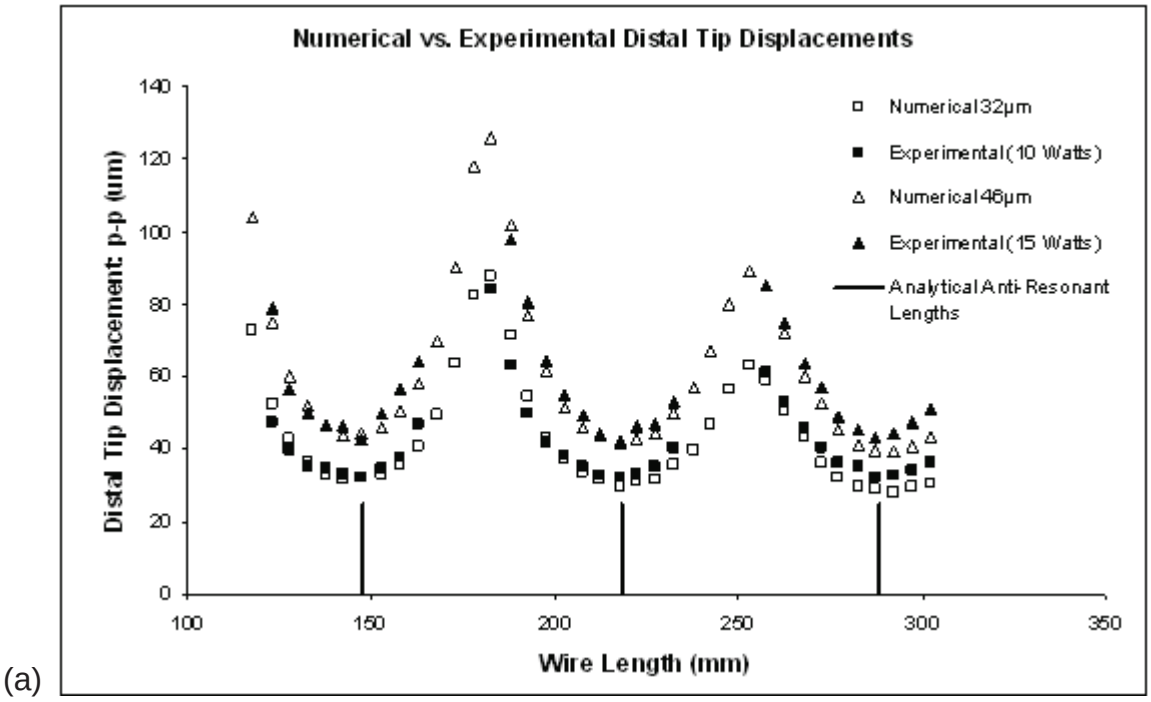
Analytical solutions related to acoustic streaming phenomena that occur around a vibrating sphere are available in the literature.[38,39] These primarily refer to the boundary layer thickness between the inner oscillatory motion and the outer acoustic streaming layer, as shown in Figure 4, and will not be discussed here in any great detail. There is also, however, a solution for the velocity field of the outer streaming layer given by Lee and Wang,[39] and this outer streaming layer can be observed experimentally.[38]

4. Recent Progress in Numerical Modeling and Design

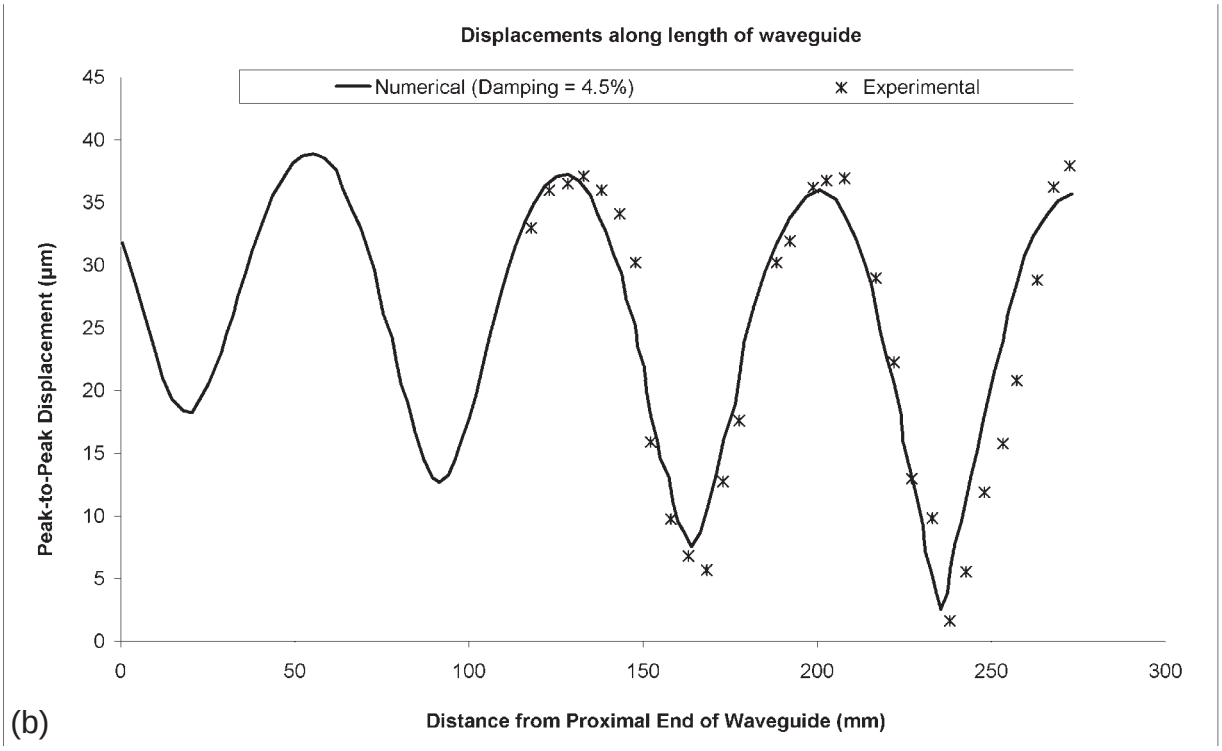
The use of computational modeling is well reported for focused ultrasound applications, predicting the pressure amplitudes, focused field effects, and thermal effects in fluid and surrounding biological tissues. However, the frequencies modeled are generally in the megahertz range, with acoustic pressure predictions up to 6 MPa. For high acoustic pressure fluctuations, non-linear effects may become significant.[54]

No computational models for predicting lesion or damage formation are yet availablefor high-power, kilohertz-frequency ultrasound wire waveguide procedures, despite the recent emergence of such medical devices.

Gavin et al. originally contributed a finite-element model of a slender rod subjected to high-frequency mechanical displacements at the proximal end, and presented predictions of the effect of wire length on distal output (peak-peak displacements).[55] The model predictions were validated by comparison with experimental measurements made using an optical microscope. Figure 7a shows experimental measurements and finite-element analysis predictions of the distal output peak-to-peak displacement for a range of wire waveguide lengths. This graph reveals the critical dependence of output on wire length, and also illustrates the usefulness of the numerical model for waveguide design. Figure 7b shows the peak-to-peak displacement amplitude at locations along the length of the wire waveguide, also well predicted by the model.



(a)



(b)

FIGURE 7. (a) Wire waveguide distal displacement amplitudes (From Gavin et al.[55]; reproduced with kind permission of Elsevier). (b) Wire waveguide internal displacement amplitudes (From Gavin et al.[56]; reproduced with kind permission of the ASME.)

These investigators also considered the effect of the damping characteristics of the wire, and showed that the nickel-titanium (NiTi) wire waveguides exhibited significant damping over waveguides approximately 300 mm in length, consistent with a damping factor of 4.5%. In later work, an acoustic fluid structure interaction finite-element model of the waveguide and surrounding fluid was developed, which accounted for the interaction of the wire distal tip with the surrounding fluid and also provided predictions of acoustic pressure in the fluid at the distal tip

Gavin et al. presented an acoustic fluid-structure model of a therapeutic angioplasty device that can predict the pressure amplitudes in the fluid field surrounding a vibrating waveguide tip.[55-57] The coupled fluid structure acoustic model assembles the following matrices of equations, where the M and K matrices represent mass and stiffness matrices for the solid (subscript S) and the fluid (subscript F), respectively.

$$\begin{bmatrix} M_S & 0 \\ \rho_0 R^T & M_F \end{bmatrix} \begin{pmatrix} \ddot{\mathbf{u}} \\ \ddot{\mathbf{P}} \end{pmatrix} + \begin{bmatrix} K_S & -R \\ 0 & K_F \end{bmatrix} \begin{pmatrix} \mathbf{u} \\ \mathbf{P} \end{pmatrix} = \begin{pmatrix} \mathbf{F}_S \\ \mathbf{P} \end{pmatrix} \qquad (7)$$

Where $[M_S]$ and $[M_F]$ are the mass matrices for the solid and fluid components, $[K_S]$ and $[K_F]$ are the stiffness matrices for the solid and fluid components; $\ddot{\mathbf{u}}$ and $\mathbf{u}$ are the nodal acceleration vector and nodal displacement vector, respectively; $\ddot{\mathbf{P}}$ and $\mathbf{P}$ are the second derivative of nodal pressure and the nodal pressure, respectively; $\mathbf{F}_S$ is the applied load vector; $[R]$ is a coupling matrix that represents effective area associated with each node, and R^T is its transpose; and ρ_0 is the fluid density. The model is capable of predicting the effect of waveguide geometry changes, such as wire length, on the instrument's resonant response and on the transmission of acoustic energy to the surrounding fluids.

Figure 8a shows the predicted acoustic pressure field around the distal tip of a wire waveguide with a 1.0-mm[en]diameter ball tip, and Figure 8b shows a comparison with published results for that configuration. This supports the validity of the finite-element model for predicting acoustic pressure fields generated by the waveguide, which is relevant to the determination of the onset of cavitation. Figure 9a illustrates visually the fluid (water) response at various peak-to-peak displacements, showing cavitation

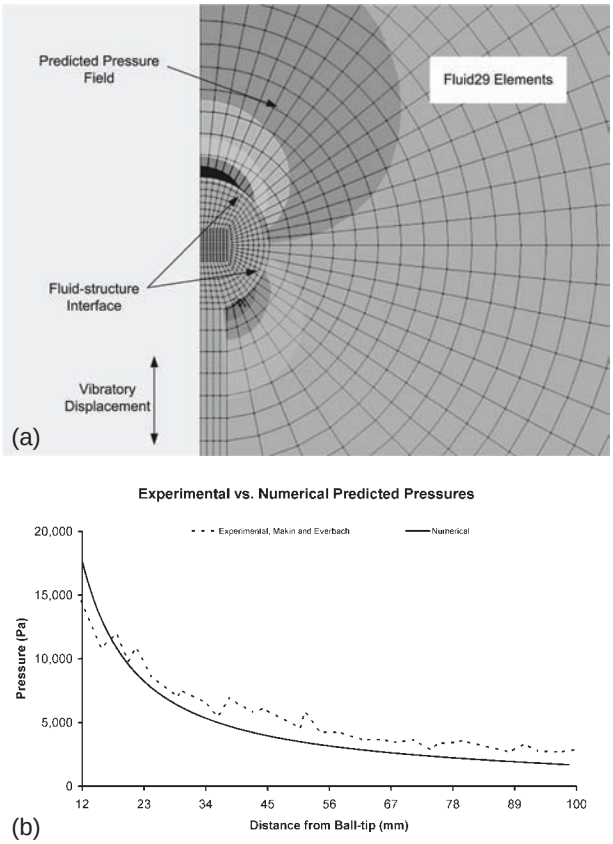


FIGURE 8. (a) Predicted pressure amplitude field around the distal tip of the wire waveguide with a 1.0-mm spherical tip. (b) Comparison of predicted pressure amplitudes and experimental results published by Makin and Everbach[44] (From Gavin et al.[56]; reproduced with kind permission of the ASME.)

developing from around the 63-μm peak-to-peak displacement. Based on a cavitation threshold of 2.5 W/cm^2, this agrees with the predictions of the numerical model. Figure 9b shows a visualization of the acoustic streaming phenomenon. The present form of the numerical model does not have the capability to predict this effect.

V. DEVICE PERFORMANCE CHARACTERIZATION AND VALIDATION

A. Transmission of Ultrasound Delivered via Wire Waveguide

While the theoretical analysis discussed provides valuable insight into the mechanics of ultrasound

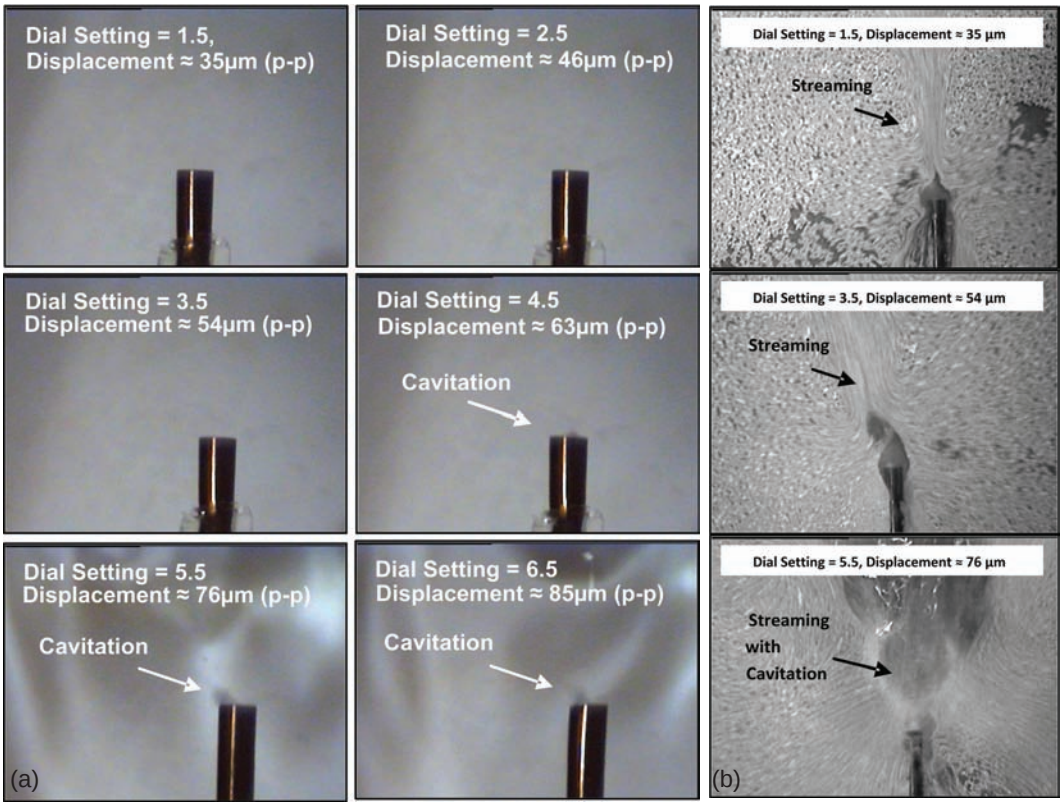


FIGURE 9. (a) Images of the distal tip of the 1.0-mm[en]diameter wire waveguide at ambient temperature for various input power dial settings between 1.5 and 6.5.[56] (b) Acoustic streaming at distal tip (1.0-mm diameter).[56] (From Gavin et al.[56]; reproduced with kind permission of the ASME.)

transmission via wire waveguide and through surrounding fluids, they are of limited use in analyzing complex geometries and conditions. The use of the finite-element method may be of benefit in assessing effects such as the inclusion of specific distal-tip geometries on the waveguide, material damping in the waveguide, and ultimately the modeling of various distal-tip shapes and their effects on the surrounding acoustic pressure field developed, with consequences for the ablation mechanisms

Further, the method may be applied to the coupling of the waveguide and surrounding fluid and modeling the acoustic fluid-structure behavior of the system. No previous modeling of therapeutic ultrasound in small-diameter wire waveguides or coupling with a distal-tip fluid appears in the literature. This technique, however, has been used in the modeling of ultrasonic acoustic horns.[58]

B. Mechanical Performance Evaluation

Gavin et al.[55] measured distal-tip displacements for short, 1.0-mm[en]diameter wire waveguides optically using a microscope with a magnification of 40 (Figure 4). This was carried out over various power inputs setting of the generator.

Gavin et al.[59] examined the 1-mm[en]diameter wire waveguides over multiple lengths, and each length was tested at various power level settings. It was found that as the power input is increased then the output peak-to-peak displacement is also increased. This is not true for all wire waveguide lengths, because some lengths result in poor or no transmission. The study suggests that the resonant length of the wires affects the transmission. Wire waveguide lengths were examined between 118 and 303 mm at intervals of 5 mm, and again using the 1-mm[en]diameter wire waveguides at a constant power setting

input. The results showed that the wire waveguide has specific resonant and anti-resonant lengths, and also that stable ultrasound transmission was attained at anti-resonant lengths. The performance decreases as the wire length approaches resonant lengths for the same power input.

C. Prototypes and Clinical Devices

Ariani et al. used an ultrasound waveguide to disrupt human thrombus in vitro and in vivo.[36] While no exact mode of disruption was determined, their results show an inverse relationship between acoustic horn distal-tip power and time to dissolve the clots; higher powers dissolved the clots in a shorter time period. Other investigators have published similar results in the treatment of thrombus both in vivo and in vitro.[23,60]

In addition to this experimental work conducted on thrombus, other investigators have tested the effect of the ultrasound delivery on atherosclerotic lesions. Siegel et al.[24] experimentally tested a 19.5-kHz system on 50 lesions of which 34% were calcified. The results showed that ultrasound delivered via a wire waveguide could recanalize the arteries affected by the lesions, and that the calcified arteries treated with therapeutic ultrasound could subsequently be dilated at lower pressures than, before suggesting an overall increase in the distensibility of the lesion.

This effect is further supported by work conducted by Demer et al.,[30] who showed that the distensibility of calcified lesions increased following ultrasound delivery. This was assessed by pressure-volume inflation curves from a standard dilation catheter, as shown in Figure 10. Following the administration of ultrasound, the inflation curve shifted to the right, indicating increased distensibility and allowed standard balloon dilation at lower pressures.

This increase in distensibility is thought to be further increased during intervention and immediately afterwards by a phenomenon called vasodilation, the relaxation of the fibers in arteries in the vicinity of low-frequency ultrasound.[35] This is thought to be a biological response and is beyond the scope of this work, but its effects are worth noting.

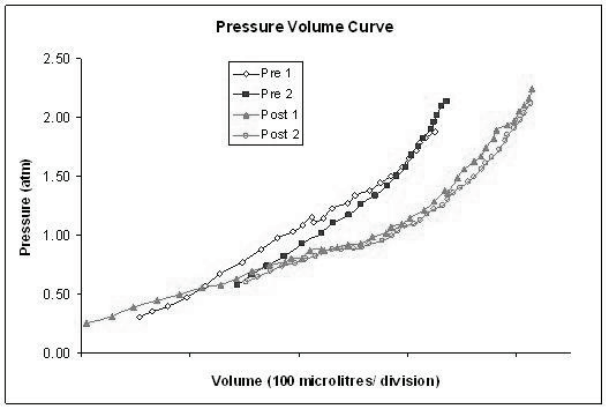


FIGURE 10. Example of pressure-volume curves obtained before (left curves) and after (right curves) ultrasound energy application. The rightward shift indicates increased distensibility. (Data digitized from Demer et al.[30] using xyExtract© digitizing software).

D. Patent Literature

This section is not intended to cover all patents, but rather to give an oversight of the main technical design issues and to show technical progression and suggested solutions. Table 3 gives an overview of significant patents in this area.The overall external electromechanical systems are similar, resulting from earlier technology requiring an ultrasonic generator, transducer (piezoelectric or magnetostrictive), and acoustic horn attached to some small-diameter waveguide effectors for insertion into the body. In addition, many other design concepts are also patented, from pulse-mode operation to user interfaces; these are considered peripheral and will not be covered in detail.

Early patents such as that obtained by Balamuth[61] describe a device that utilizes ultrasonic energy for the removal of warts, tumors, and skin cancer but makes no mention of operating within blood vessels. Kuris[62] obtained one of the earliest patents to disclose the use of vibratory displacements transmitted through a catheter inserted into a blood vessel that is surgically exposed. One of the first patents on the road to current approaches was by Boyd,[63] who describes the use of a semi-flexible ultrasonic vibratory catheter for the treatment of atheromas including coronary atherosclerosis. In this invention, the hollow catheter itself vibrates and has a sharpened tip to facilitate the removal of tissue.

TABLE 3. Patent Literature

Year	Patent Number	Inventor	Assignee
1969	US 3,433,226	Boyd	Aeroprojects Inc.
1970	US 3,526,219	Balamuth	Ultrasonic Systems
1971	US 3,565,062	Kuris	Ultrasonic Systems
1989	US 4,870,953	DonMicheal, Siegel, and DeCastro	--
1989	WO 89/06515	Bernstein and Rosenschein	--
1995	US 5,397,301	Pflueger, Nita, Siegel, Bacich, Bond, and DeCastro	Baxter International Inc.
1999	US 5,971,949	Levin, Saltonstall, Nguyen, and Rosenschein	Angiosonics Inc.
2006	EP 1649,817 A2	Nita	FlowCardia Inc.
2009	US 2009/0216246	Nita, Sarge. and Spano	FlowCardia Inc.

Limitations of these early patents include the need for open surgery, stiff transmission probes not suited to minimally invasive access, associated risks of arterial perforation, and other complications. To overcome these, devices needed to be designed that could deliver ultrasonic vibrations via small-diameter, flexible waveguides (lengths approximately 1.6 m and diameters less than 1 mm) similar to those used in other minimally technologies, such as balloon dilation catheters.

As a result, the focus of the detailed technical design, solutions, and subsequent patents in the area of minimally invasive ultrasound catheter systems for the ablation and disruption of arterial plaques, are in three critical areas:

1. Wire waveguide design and material choice,
2. The method by which the waveguide is connected to the acoustic horn, and
3. The wire waveguide distal-tip configuration.

Two key patents filed in 1989 describe ultrasonic angioplasty systems incorporating flexible waveguides, detailed connection methods, and peripheral tip designs. Bernstein and Rosenschein[64] describe an ultrasonic angioplasty waveguide system using an aluminum-based alloy (AL-7075 etc.) for the waveguide material to reduce attenuation of the ultrasonic energy over long working lengths. They state that the distal tip should preferably be "flat." Their method of connecting the wire waveguide to the acoustic horn uses a connector that applies a gripping force to the waveguide by two radially opposing grub screws. Shrink-fitting the connector and waveguide using heat is also mentioned.

DonMichael et al.[65] describe an apparatus that can induce both longitudinal and transverse motions to the wire waveguide. They propose a cobalt-nickel alloy as a suitable waveguide material. It is also suggested that a bulbous (enlarged) distal tip efficiently transfers the ultrasonic energy while reducing the risk of "perforating the artery." Suggested connection methods to the acoustic horn include a "vibration fitting," but no specific details are provided. After 1991, many patents became increasingly more detailed.

In 1995, a patent by Pflueger et al.[66,67] including some of the same inventors from above (Siegel and DeCastro) proposed an ultrasonic device comprising an elongate ultrasound transmission member (including tapered waveguides) formed of one or more super-elastic alloys such as NiTi (51% Ni by atomic percent). A range of wire waveguide distal-tip designs are disclosed, including surface finishes with grooves and dimples to enhance material disruption and cavitation. A ferrule grip was suggested as a means of connecting the waveguide to the acoustic horn. Also mentioned was using the wire waveguide in an "over-the-wire"' configuration similar to bal-

loon angioplasty and more common interventional methods. In 1999, Levin et al.[68] described a system with a waveguide consisting of a series of transitions, multiple wires in parallel, and ideal transition locations based on the standing wave setup in the waveguide. Waveguides consisting of multiple materials transitioning along the length, such as aluminum and titanium alloys, were included.

Nita and Sarge[67] disclose an ultrasound system with wire waveguides made from a material that included nickel (50.5–51.5 at.) having tapered sections, and in their embodiment they suggest NiTi superelastic alloys. A number of connection methods and distal-tip configurations are disclosed, including a balloon to facilitate off-center positioning of the wire waveguide distal tip.

In 2009, Nita et al.[69] described an ultrasound catheter device for the disruption of occlusions of blood vessels. They describe a wire waveguide catheter system including irrigation ports for cooling of the waveguide, an over-the-wire configuration, and a detailed connection system to eliminate lateral movements of the waveguide near the connection point, all with the goal of reducing heat in the waveguide and increasing its longevity.

VI. FUTURE DIRECTIONS

Much work is still required to further understand this form of ultrasound energy delivery and the subsequent affects this has on plaque material and biological response. Regarding the fundamental principles of the process, there is a lack of clarity surrounding the relative importance of the various proposed disruption mechanisms. Work is needed to understand the conditions required to cause cavitation in blood in vivo and to determine its importance in plaque disruption. There is also a need to isolate tissue or plaque effects due principally to acoustic streaming, and to evaluate the role that this mechanism can play. Of greatest importance perhaps is the need to develop an understanding of the fracture behavior of plaque caps and lipid pools under contact with a surface vibrating at low ultrasonic frequencies.

There is a need to develop a greater understanding of the mechanics of wave transmission through long, flexible wire waveguides with complex material characteristics (as exemplified by NiTi superelastic shape memory alloys). Significant damping of wave amplitude has been observed in straight NiTi wires at low ultrasonic frequencies, and the effect of tortuous pathways that necessitate significant bending require much greater investigation with a view to controlling losses. It is essential that the distal-tip vibrations can be precisely and consistently controlled by the cardiologist, and are not unduly influenced by variable factors associated with individual procedures.

The fatigue of superelastic metals under cyclic strain at low ultrasonic frequencies has not been extensively studied, even though the reliability of these devices depends on these properties. Designs must, in any case, be configured such that wire fatigue failures (more likely at locations of mechanical connection, in the experience of the authors) do not result in being unable to retrieve all of the distal elements of the device.

VII. SUMMARY

Calcified and totally occluded arteries pose significant complications during standard balloon angioplasty interventions and stent implantations. The use of therapeutic ultrasound transmitted via wire waveguides has been shown in initial clinical testing to be capable of disrupting thrombus and calcified and fibrous plaque without significant damage to healthy arterial tissue. Potential end clinical benefits in the use of therapeutic ultrasound delivered via small-diameter[en]wire waveguides includes increasing the distensibility of calcified plaques with the potential of reducing barotraumas and restenosis rates. This technology has also the potential in crossing CTOs, allowing for standard balloon angioplasty and stent implantation.

Little detailed analysis is provided on the design aspects of these devices or how ultrasonic displacements are delivered to the distal tips of these small-diameter[en]wire waveguides at the frequencies and amplitudes required and the disruption mechanisms these cause. With the launch of products to the U.S. and European markets potentially imminent, sophisticated models for the design and optimization of these devices are urgently required. This challenge

demands a combination of experimental and clinical investigation and computational modeling. A validated finite-element model of the mechanical behavior of a small-diameter[en]wire waveguides and their interaction with surrounding fluids is necessary. Any such model would prove highly beneficial in determining the crucial wire waveguide distal-tip displacements, and should be capable of predicting when the highly disruptive cavitation events occur.

REFERENCES

1. Haust MD. The natural history of human atherosclerotic lesions. In: Moore S, ed. Vascular injury and atherosclerosis. New York: Marcel Dekker; 1981. p. 1–23.
2. Woolf N. Pathology of Atherosclerosis. 1st ed. London: Butterworth Scientific; 1982.
3. Ross R. The pathogenesis of atherosclerosis-an update. N Engl J Med. 1986 Feb 20;314(8):488–500.
4. Bouissou M, Pierragi T, Julian M. Progression, topological aspects and regression of atherosclerosis. In: Camilleri JP, Berry CL, Fiessinger JN, Bariety J, editors. Diseases of the arterial wall. Berlin: Springer-Verlag; 1989.
5. Salunke NV, Topoleski LDT. Biomechanics of atherosclerotic plaque. Crit Rev Biomed Eng. 1997;25(3):243–85.
6. Stary HC. The histological classification of atherosclerotic lesions in human coronary arteries. In: Fuster V, Ross R, Topol EJ, editors. Atherosclerosis and coronary artery disease. Philadaelphia: Lippincott Williams & Wilkins; 1996. p. 463–74.
7. Loree HM, Grodzinsky AJ, Park SY, Gibson LJ, Lee RT. Static circumferential tangential modulus of human atherosclerotic tissue. J.Biomech. 1994 Feb;27(2):195–204.
8. Topoleski LD, Salunke NV. Mechanical behavior of calcified plaques: a summary of compression and stress-relaxation experiments. Z Kardiol. 2000;89 Suppl 2:85–91.
9. Ng W, Chen WH, Lee PY, Lau CP. Initial experience and safety in the treatment of chronic total coronary occlusions with a new optical coherent reflectometry-guided radiofrequency ablation guidewire. Am J Cardiol. 2003 Sep 15;92(6):732–4.
10. Aziz S, Ramsdale DR. Chronic total occlusions--a stiff challenge requiring a major breakthrough: is there light at the end of the tunnel? Heart. 2005 Jun;91 Suppl 3:iii42–8.
11. Katsuragawa M, Fujiwara H, Miyamae M, Sasayama S. Histologic studies in percutaneous transluminal coronary angioplasty for chronic total occlusion: comparison of tapering and abrupt types of occlusion and short and long occluded segments. J Am Coll Cardiol. 1993 Mar 1;21(3):604–11.
12. Stone GW, Reifart NJ, Moussa I, Hoye A, Cox DA, Colombo A, Baim DS, Teirstein PS, Strauss BH, Selmon M, Mintz GS, Katoh O, Mitsudo K, Suzuki T, Tamai H, Grube E, Cannon LA, Kandzari DE, Reisman M, Schwartz RS, Bailey S, Dangas G, Mehran R, Abizaid A, Moses JW, Leon MB, Serruys PW. Percutaneous recanalization of chronically occluded coronary arteries: a consensus document: part II. Circulation. 2005 Oct 18;112(16):2530–7.
13. Olivari Z, Rubartelli P, Piscione F, Ettori F, Fontanelli A, Salemme L, Giachero C, Di Mario C, Gabrielli G, Spedicato L, Bedogni F; TOAST-GISE Investigators. Immediate results and one-year clinical outcome after percutaneous coronary interventions in chronic total occlusions: data from a multicenter, prospective, observational study (TOAST-GISE). J Am Coll Cardiol. 2003 May 21;41(10):1672–8.
14. Stone GW, Kandzari DE, Mehran R, Colombo A, Schwartz RS, Bailey S, Moussa I, Teirstein PS, Dangas G, Baim DS, Selmon M, Strauss BH, Tamai H, Suzuki T, Mitsudo K, Katoh O, Cox DA, Hoye A, Mintz GS, Grube E, Cannon LA, Reifart NJ, Reisman M, Abizaid A, Moses JW, Leon MB, Serruys PW. Percutaneous recanalization of chronically occluded coronary arteries: a consensus document: part I. Circulation. 2005 Oct 11;112(15):2364–72.
15. Stone GW, Colombo A, Teirstein PS, Moses JW, Leon MB, Reifart NJ, Mintz GS, Hoye A, Cox DA, Baim DS, Strauss BH, Selmon M, Moussa I, Suzuki T, Tamai H, Katoh O, Mitsudo K, Grube E, Cannon LA, Kandzari DE, Reisman M, Schwartz RS, Bailey S, Dangas G, Mehran R, Abizaid A, Serruys PW. Percutaneous recanalization of chronically occluded coronary arteries: procedural techniques, devices, and results. Catheter Cardiovasc Interv. 2005 Oct;66(2):217–36.
16. Folland ED. Balloon angioplasty. In: Topol EJ, Serruys, PW, editors. Current review of interventional cardiology. Philadelphia: Current Medicine; 1994.
17. Oh S, Kleinberger M, McElhaney JH. A finite element analysis of balloon angioplasty. In: Advances in Bioengineering: Anaheim, CA: ASME; 1992. p. 269–72.
18. Oh S, Kleinberger M, McElhaney JH. Finite-element analysis of balloon angioplasty. Med Biol Eng Comput. 1994 Jul;32(4 Suppl):S108–14.
19. Godino C, Sharp AS, Carlino M, Colombo A. Crossing CTOs-the tips, tricks, and specialist kit that can mean the difference between success and failure. Catheter Cardiovasc Interv. 2009 Dec 1;74(7):1019–46.
20. Lefevre T, Louvard Y, Loubeyre C, Dumas P, Piechaud JF, Krol M, Benslimane A, Premchand RK, Morice MC. A randomized study comparing two guidewire strategies for angioplasty of chronic total coronary occlusion. Am J Cardiol. 2000 May 1;85(9):1144–7.
21. Siegel RJ, Fishbein MC, Forrester J, Moore K, DeCastro E, Daykhovsky L, DonMichael TA, Ultrasonic plaque ablation. A new method for recanalization of partially or totally occluded arteries. Circulation, 1988 Dec;78(6):1443–8.
22. Rosenschein U, Bernstein JJ, DiSegni E, Kaplinsky E, Bernheim J, Rozenzsajn LA. Experimental ultrasonic angioplasty: disruption of atherosclerotic plaques and thrombi in vitro and arterial recanalization in vivo. J Am Coll

Cardiol. 1990 Mar 1;15(3):711–7.
23. Rosenschein U, Rozenszajn LA, Kraus L, Marboe CC, Watkins JF, Rose EA, David D, Cannon PJ, Weinstein JS. Ultrasonic angioplasty in totally occluded peripheral arteries. Initial clinical, histological, and angiographic results. Circulation. 1991 Jun;83(6):1976–86.
24. Siegel RJ, Gaines P, Crew JR, Cumberland DC. Clinical trial of percutaneous peripheral ultrasound angioplasty. J Am Coll Cardiol. 1993 Aug;22(2):480–8.
25. Siegel RJ, Gunn J, Ahsan A, Fishbein MC, Bowes RJ, Oakley D, Wales C, Steffen W, Campbell S, Nita H. Use of therapeutic ultrasound in percutaneous coronary angioplasty. Experimental in vitro studies and initial clinical experience. Circulation. 1994 Apr;89(4):1587–92.
26. Melzi G, Cosgrave J, Biondi-Zoccai GL, Airoldi F, Michev I, Chieffo A, Sangiorgi GM, Montorfano M, Carlino M, Colombo A. A novel approach to chronic total occlusions: the crosser system. Catheter Cardiovasc Interv. 2006 Jul;68(1):29–35.
27. Grube E, Sütsch G, Lim VY, Buellesfeld L, Iakovou I, Vitrella G, Colombo A. High-frequency mechanical vibration to recanalize chronic total occlusions after failure to cross with conventional guidewires. J Invasive Cardiol. 2006 Mar;18(3):85–91.
28. Yock PG, Fitzgerald PJ. Catheter-based ultrasound thrombolysis. Circulation. 1997 Mar 18;95(6):1360–2.
29. Atar S, Luo H, Nagai T, Siegel RJ. Ultrasonic thrombolysis: catheter-delivered and transcutaneous applications. Eur J Ultrasound. 1999 Mar;9(1):39–54.
30. Demer LL, Ariani M, Siegel RJ. High intensity ultrasound increases distensibility of calcific atherosclerotic arteries. J Am Coll Cardiol. 1991 Nov 1;18(5):1259–62.
31. Sobbe A, Stumpff U, Trubenstein G, Figge H, Kozuchek W. Die Ultraschall-Auflosung von Thromben [article in German]. Klin Wochenschr. 1974;52:1117–21.
32. Wevers M, Lafaut JP, Baert L, Chilibon I. Low-frequency ultrasonic piezoceramic sandwich transducer. Sensors and Actuators A: Physical. 2005;122(2):284–9.
33. Peshkovsky SL, Peshkovsky AS. Matching a transducer to water at cavitation: acoustic horn design principles. Ultrasonics: Sonochemistry. 2007 Mar;14(3):314–22.
34. Yang JJ, Fang ZD, Wei BY, Deng XZ. Theoretical explanation of the 'local resonance' in stepped acoustic horn based on Four-End Network method. Journal of Materials Processing Technology. 2009;209(6):3106–10.
35. Fischell TA, Abbas MA, Grant GW, Siegel RJ. Ultrasonic energy. Effects on vascular function and integrity. Circulation. 1991 Oct;84(4):1783–95.
36. Ariani M, Fishbein MC, Chae JS, Sadeghi H, Michael AD, Dubin SB, Siegel RJ. Dissolution of peripheral arterial thrombi by ultrasound. Circulation. 1991 Oct;84(4):1680–8.
37. Burdic WS. Underwater acoustic system analysis. 2nd ed. Upper Saddle River, NJ: Prentice Hall; 1991.
38. Nyborg WL. Basic physics of low frequency therapeutic ultrasound. In: Siegel RJ, editor. Ultrasound angioplasty: developments in cardiovascular medicine. Berlin: Kluwer Academic Publishers; 1996.
39. Lee CP, Wang TG. Outer acoustic streaming. J Acoust Soc Am. 1990;40:1363–70.
40. Morse PM. Vibration and sound. New York: Acoustical Society of America; 1981. p. 311–26.
41. Thoe TB, Aspiwall DK, Wise MLH. Review of ultrasonic machining. International Journal of Machine Tools and Manufacture. 1998;38(4):239–55.
42. Ernst A, Schenk EA, Woodlock TJ, Alliger H, Gottlieb S, Child SZ, Meltzer RS. Feasibility of recanalization of human coronary arteries using high-intensity ultrasound. Am J Cardiol. 1994 Jan 15;73(2):126–32.
43. Miller DL. A review of the ultrasonic bioeffects of microsonation, gas-body activation, and related cavitation-like phenomena. Ultrasound Med Biol. 1987 Aug;13(8):443–70.
44. Makin IR, Everbach EC. Measurement of pressure and assessment of cavitation for a 22.5-kHz intra-arterial angioplasty device. J Acoust Soc Am. 1996 Sep;100(3):1855–64.
45. Porges G. Applied acoustics. London: Edward Arnold; 1977.
46. Herickhoff CD, Light ED, Bing KF, Mukundan S, Grant GA, Wolf PD, Smith SW. Dual-mode intracranial catheter integrating 3D ultrasound imaging and hyperthermia for neuro-oncology: feasibility study. Ultrason Imaging. 2009 Apr;31(2):81–100.
47. Herickhoff CD, Light ED, Wolf PD, Smith SW, Grant GA, Britz GW. Dual-mode intracranial catheters for minimally-invasive neuro-oncology. In: Proceedings of the 2009 IEEE International Ultrasonics Symposium. Piscataway, NJ: IEEE; 2009. p. 1012–25.
48. Lafon C, de Lima DM, Theillere Y, Prat F, Chapelon JY, Cathignol D. Optimizing the shape of ultrasound transducers for interstitial thermal ablation. Med Phys. 2002;29(3):290–7.
49. Lafon C, Melodelima D, Salomir R, Chapelon JY. Interstitial devices for minimally invasive thermal ablation by high intensity ultrasound. Int J Hyperthermia. 2007 Mar;23(2):153–63.
50. Parikh S, Motarjeme A, McNamara T, Raabe R, Hagspiel K, Benenati JF, Sterling K, Comerota A. Ultrasound-accelerated thrombolysis for the treatment of deep vein thrombosis: initial clinical experience. J Vasc Interv Radiol. 2008 Apr;19(4):521–8.
51. Smikahl J, Yeung D, Wang S, Semba CP. Alteplase stability and bioactivity after low-power ultrasonic energy delivery with the OmniSonics resolution system. J Vasc Interv Radiol. 2005 Mar;16(3):385–9.
52. Steidel RF Jr. An introduction to mechanical vibrations. 3rd ed. New York: Wiley; 1989.
53. Perkins JP. Power ultrasonic equipment. Paper presented at the Sonochemistry Symposium, Annual Chemical Congress; 1986 Apr 8–11; Warwick University, UK.
54. Wojcik G, Mould J, Lizzi F, Abboud N, Ostromogilsky M, Vaughan D. Non-linear modelling of therapeutic ul-

trasound. In: Proceedings of the 1995 IEEE International Ultrasonics Symposium. Piscataway, NJ: IEEE; 1996; p. 1617–22.
55. Gavin GP, McGuinness GB, Dolan F, Hashmi MS. Performance characteristics of a therapeutic ultrasound wire waveguide apparatus. International Journal of Mechanical Sciences. 2007;49:298–305.
56. Gavin GP, Dolan F, Hashmi MSJ, McGuinness GB. A coupled fluid-structure model of a therapeutic ultrasound angioplasty wire waveguide. Journal of Medical Devices. 2007;1:254–63.
57. Wylie MP, McGuinness GB, Gavin GP. A linear finite element acoustic fluid-structure model of ultrasonic angioplasty in vivo. International Journal for Numerical Methods in Biomedical Engineering. 2010 July 2010;26(7):828–42.
58. Cunningham PM. Use of the finite element method in ultrasonic applications. Paper presented at the Ultrasonic Industry Association Symposium; 2000 June 11–14; Columbus, OH.
59. Gavin GP. Experimental and numerical investigation of therapeutic ultrasound angioplasty [PhD dissertation]. Dublin, Ireland: Dublin City University; 2005.
60. Steffen W, Fishbein MC, Luo H, Lee DY, Nita H, Cumberland DC, Tabak SW, Carbonne M, Maurer G, Siegel RJ. High intensity, low frequency catheter-delivered ultrasound dissolution of occlusive coronary artery thrombi: an in vitro and in vivo study. J Am Coll Cardiol. 1994 Nov 15;24(6):1571–9.
61. Balamuth L, inventor. Ultrasonic Systems Inc., assignee. Method and apparatus for ultrasonically removing tissue from a biological organism. United States patent US 3,526,219. 1970 September 1.
62. Kuris A, inventor. Ultrasonic Systems Inc., assignee. Ultrasonic method and apparatus for removing cholesterol and other deposits from blood vessels and the like. United States patent US 3,565,062. 1971 February 23.
63. Boyd CA, inventor. Aeroprojects Incorporated P, assignee. Vibratory catheterisation apparatus and method for using. United States patent US 3,433,226. 1969 March 18.
64. Bernstein J, Rosenschein U, inventors. System for angioplasty and ultrasonic contrast imaging. World Intellectual Property Association patent WO 89/006515. 1989 July 27.
65. DonMichael TA, Siegel RJ, Decastro EA, inventors. Intravascular ultrasonic catheter/probe and method for treating intravascular blockage. United States patent US 4,870,953. 1989 October 3.
66. Pflueger R, Nita H, Bacich S, Siegel R, Bond G, Decastro E, inventors. Ultrasonic angioplasty device incorporating an ultrasound transmission member made at least partially from a superelastic metal alloy. United States patent US 5,397,301. 1995 March 14.
67. Nita H, Sarge J, inventors. Device and method for supporting placement of a therapeutic device in a blood vessel. United States patent US 2005/0124877 A1. 2005 June 9.
68. Levin PS, Saltonstall J, Nguyen L, Rosenschein U, inventors. Ultrasound transmission apparatus and method of using same. United States patent US 5,971,949. 1999 October 26.
69. Nita H, Sarge J, Spano R, inventors. Ultrasound catheter devices and methods. United States patent US 2009/0216246. 2009 August 27.

Critical Reviews™ in Biomedical Engineering, 38(6):533–577 (2010)

Mathematical Foundations of Biomechanics

Peter F. Niederer*

*Address all correspondence to Peter F. Niederer, Prof. em. Biomedical Engineering, Swiss Federal Institute of Technology (ETH), Gloriastrasse 35, CH-8092 Zurich, Switzerland; Tel.: +41 44 256 4568; Fax: +41 44 256 1193; niederer@biomed.ee.ethz.ch.

ABSTRACT: The aim of biomechanics is the analysis of the structure and function of humans, animals, and plants by means of the methods of mechanics. Its foundations are in particular embedded in mathematics, physics, and informatics. Due to the inherent multidisciplinary character deriving from its aim, biomechanics has numerous connections and overlapping areas with biology, biochemistry, physiology, and pathophysiology, along with clinical medicine, so its range is enormously wide. This treatise is mainly meant to serve as an introduction and overview for readers and students who intend to acquire a basic understanding of the mathematical principles and mechanics that constitute the foundation of biomechanics; accordingly, its contents are limited to basic theoretical principles of general validity and long-range significance. Selected examples are included that are representative for the problems treated in biomechanics. Although ultimate mathematical generality is not in the foreground, an attempt is made to derive the theory from basic principles. A concise and systematic formulation is thereby intended with the aim that the reader is provided with a working knowledge. It is assumed that he or she is familiar with the principles of calculus, vector analysis, and linear algebra.

KEY WORDS: biomechanics; modeling; mechanobiology; biomathematics

I. INTRODUCTION

A. General Remarks

We are exposed to a great variety of forces and mechanical loading situations throughout our lives. In addition to forces deriving from ubiquitous and penetrating fields such as gravity or interactions originating from electromagnetism, which are non-contact in nature and as such effective over distances, there are numerous external loads acting on all living systems from interactions with their surroundings. In addition, numerous forces are generated in the course of physiological processes inside the body in the different organs and tissues. Throughout evolution, all forms of life adapted themselves to mechanical exposure; in fact, a lack of sufficient mechanical loading may have adverse effects. For example, muscle mass is resorbed when not used or proper bone remodeling is impeded under conditions of insufficient loading.

The science of biomechanics is devoted to the analysis, measurement, and modeling of the structure and function of humans, animals, or plants or of parts thereof by application of the methods of mechanics. In particular, the biological effects that are observed under the various mechanical loading situations are investigated. As such, biomechanics contributes to a scientific understanding of living systems. An important aim consists of providing a basis for the application of engineering principles for the improvement of living conditions in general and for the restoration of human function in case of disease or disablement in particular.

In light of its aims, biomechanics represents a multidisciplinary field characterized by numerous connections and overlapping areas with biology, biochemistry, physiology, and pathophysiology, as well as with clinical medicine, so its range is enormously wide. It includes the application of mechanics to such distant fields as human motion, rehabilitation, surgery simulation, implant mechanics, sports, trauma, cardiovascular dynamics, airflow in the lungs, organ dynamics, cell motility, transmembrane cellular transport, and the influence of mechanical forces on gene expression. Animal- and plant-related biomechanics include the flight of birds, the swimming of fishes, and the growth of trees. A quantitative approach is thereby always in the foreground, and thus, to some extent, biomechanics can be regarded as quantitative physiology and pathophysiology. The term mechanobiology refers specifically to the transduction of mechanical into biological signals. Biomechanics—as

1040-8401/10/$35.00

a basic as well as an engineering science—therefore covers the entire range from molecular, cellular, tissue, and organ to whole body dimensions.

The extent of forces that are of interest is large: internal forces may originate from the action of molecules, contractile fibers on a cellular level, or muscles on a macroscopic scale; moreover, pressures and shear stresses may be generated by biological fluid flows or active biological transport processes including osmosis. External forces occurring in everyday life may span a virtually unlimited extent. Accordingly, the forces of interest in biomechanics typically cover a range from the piconewton to the meganewton (lower or higher forces, respectively, are hardly considered because of lack of biological effect on the lower side or complete devastation on the upper), and their effects may manifest themselves within picoseconds (molecular actions) or years (bone remodeling).

An attempt is made in this treatise to demonstrate the basic mathematical and physical principles upon which biomechanics is based. Ultimate mathematical generality is thereby not in the foreground; for example, only rectangular coordinate systems are used. In particular, a concise presentation is sought. For details, the reader is referred to the specialized literature. Examples are intended to demonstrate representative applications of the treated theoretical material. Likewise, references are limited to a selection of representative and seminal publications and textbooks and are by far not exhaustive. They should primarily reflect the wide range of subjects treated in biomechanics.

B. Historical Remarks

The science of biomechanics has a long history. Phenomena such as bird flight or the beating heart, along with the pulse that can be felt at certain locations of the human body, stimulated the interest of scientists through centuries. Already in ancient Chinese medicine, pulse-wave analysis served as a major diagnostic procedure. A number of more recent scholars who made major contributions to biomechanics as we understand it today are:

- Santorio Santorio (1561–1636) was one of the first scientists to introduce quantitative methods into biology. He also invented various medical instruments and investigated human metabolism by systematic whole-body weighing experiments (food intake/feces excretion balance).
- Galileo Galilei (1564–1642) measured pulse rates and made observations with respect to weight and size of animals. He also made important contributions to the development of microscopy.
- William Harvey (1578–1658) postulated the circulatory nature of the mammalian blood flow system after having studied the functioning of the heart. However, he could not demonstrate a connection between the arterial and venous side because microscopy did not have sufficient resolution at that time.
- Giovanni Alfonso Borelli (1608–1679), sometimes referred to as the "Father of Biomechanics" devoted much of his life to the study of animal movement and bird flight.
- Robert Boyle (1627–1691), when working with air and pumps, made observations on the lungs and associated air flow.
- Marcello Malpighi (1628–1694) discovered the capillaries and is considered as one of the founders of microscopic anatomy.
- Robert Hooke (1635–1703), Boyle's assistant, is known for Hooke's law of elasticity. He also introduced the term "cell" as the smallest unit of life.
- Stephen Hales (1677–1761) put forward the "Windkessel" notion of the arterial system (although he did not use this particular expression).
- Leonhard Euler (1707–1783) is the founder of continuum mechanics, in particular of fluid mechanics. In his *Principia pro motu sanguinis per arterias determinando* (op. posth.), he applied his theoretical principles to the mathematical modeling of blood flow.
- Jean Poiseuille (1797–1869) studied blood flow and is believed to be the first scientist to have measured blood pressure.
- Hermann von Helmholtz (1821–1894) made many contributions to physiology in general, many of them with relevance for

biomechanics (sound perception, physiologic energy balance, geometry and optics of the eye, etc.)

- Adolf Fick (1829–1901), best known for Fick's law of diffusion, devised a method to measure cardiac output (a modern variant is still used today).
- Diederick Johannes Korteweg (1848–1941), A. I. Moens (*Die Pulskurve*, Leiden 1878), Horace Lamb (1849–1934), Konrad Witzig (1864–1962), Otto Franck (1865–1944), and John R. Womersley (1907–1958) produced a large body of work relating to blood flow, arterial pulse propagation, and cardiac mechanics.
- Otto Messerer (main work in this context, *Über Elastizität und Festigkeit der menschlichen Knochen*, Wertheim München, 1880) pioneered forensic biomechanics.
- Carl Culmann (1821–1881), Georg H. von Meyer (1815–1892), Karl W. Ritter (1847–1906), and above all Julius Wolff (1836–1902, main work *Das Gesetz der Transformation der Knochen*, A. Hirschwald Berlin, 1982) investigated the connection between mechanical loading and bone and skeletal architecture and remodeling.

C. Physical and Mathematical Approximations Used in Biomechanics

From a mechanics' point of view, biological materials behave in a highly nonlinear fashion: they are anisotropic, viscoelastic, and exhibit an inhomogeneous composition ("structured continua"). Due to fiber rearrangement and interstitial fluid displacement during deformation, biological tissues also exhibit quasi-plastic properties. Various tissues also have the ability to produce forces actively themselves due to intrinsic contractile elements, in particular, muscles. For practical purposes, however, substantial and rigorous approximations are necessary to analyze systems of such complexity.

The simplest mechanical object is a point mass. However, such an approximation is hardly used in general biomechanics because it lacks any structure; only molecular models are partly based on this approximation. Rigid body approximations, in turn, have various applications, in particular for human or animal motion analysis or for a simplified preliminary investigation of an otherwise complicated process. Often, systems of rigid bodies are considered, which are connected in joints (e.g., a model for the human body) or have interactions through external forces (e.g., molecules). It should also be noted that according to the solidification principle, also in the case of a deformable body, the center of mass obeys Newton's second law of motion much as a point mass. The problem with this is that the center of mass changes its location relative to the body contour in a not-straightforward fashion when it is being deformed.

When tissue deformations, organ dynamics, cellular processes, or biofluid motions are considered, methods of continuum mechanics or hybrid approaches have to be applied. The constitutive properties of the biological materials in question are thereby modeled in various degrees of complexity.

From a mathematical point of view, an important distinction between rigid body and continuum approximations derives from the fact that rigid body systems have a finite number of degrees of freedom, whereas in continuum mechanics the degree of freedom is infinite. Rigid body systems are therefore described by ordinary and continuum mechanics by partial differential equations. The complexity of biological systems is such that computer-based approaches are widely applied. To this end, discretization in the case of continuum modeling is necessary. The most appropriate discretization scheme has to be determined depending on the problem to be solved.

D. Mathematical Formulations; Vector and Tensor Notations

Extensive use is made in this text of vectors and tensors (both written in bold italics). In practical solution procedures, vectors and tensors are represented through their components in a reference system. Particularly useful is therefore the representation in components of vectors as row or column triples or second-order tensors as matrices. In so doing, however, one has to keep in mind that one

has to distinguish mathematically between a vector or tensor as a formal element and the set of their components in a coordinate system. Thus, only rectangular coordinate systems are used, such that covariant and contravariant components need not be distinguished. In a Cartesian coordinate system, a vector, $\boldsymbol{x}$, e.g., a velocity, a force or a position vector can be decomposed according to $\boldsymbol{x} = \boldsymbol{x}_1 \boldsymbol{e}_{x_1} + \boldsymbol{x}_2 \boldsymbol{e}_{x_2} + \boldsymbol{x}_3 \boldsymbol{e}_{x_3}$ with coordinates $\boldsymbol{x}_1$, $\boldsymbol{x}_2$, $\boldsymbol{x}_3$ and orthogonal unit vectors $\boldsymbol{e}_{x_1}$, $\boldsymbol{e}_{x_2}$, $\boldsymbol{e}_{x_3}$. In components, $\boldsymbol{x}$ will be written in brackets as $(\boldsymbol{x}) = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$, or $(\boldsymbol{x})_i = \boldsymbol{x}_i$, $i = 1,2,3$. Also the notation notation $(\boldsymbol{x})^{\mathsf{T}} = (\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3)$ ($^{\mathsf{T}}$ for transposed) is also used. Likewise, a second-order tensor, $\boldsymbol{T}$, e.g., the /stress tensor, will be represented by its components as: $(\boldsymbol{T}) = \begin{pmatrix} t_{11} & t_{12} & t_{13} \\ t_{21} & t_{22} & t_{23} \\ t_{31} & t_{32} & t_{33} \end{pmatrix}$.

A scalar product between vectors or tensors is written as (..., ...) whereby the notation with brackets is used to denote that the expression represents a formulation in components, i.e., it can be performed as a matrix multiplication. It involves transposition of the first term and subsequent multiplication, for example:

$$(\boldsymbol{x}, \boldsymbol{x}) = (\boldsymbol{x})^{\mathsf{T}} (\boldsymbol{x}) = (x_1, x_2, x_3) \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = x_1^2 + x_2^2 + x_3^2.$$

Vector products are written using the symbol (... $\times$ ...). Such products will also appear only in component formulation. The dot, $\cdot$, in turn, will be inserted occasionally to indicate a multiplication of vectors and tensors.

Many problems in biomechanics are such that modeling in curvilinear coordinates is more appropriate than in a Cartesian system; for example, blood vessels, urinary ducts, and airways in the lungs. Cylindrical or spherical coordinate systems that are curvilinear, yet rectangular, are applied in such cases. An analogous notation can be used as above; nevertheless, when calculus is performed, the curvilinearity has to be taken into account.

To this end, we consider a transformation of a set of Cartesian coordinates $\boldsymbol{x}_1$, $\boldsymbol{x}_2$, $\boldsymbol{x}_3$ into a curvilinear, orthogonal system ξ_1, ξ_2, ξ_3 by way of the functions

$$\begin{aligned} x_1 &= x_1(\xi_1, \xi_2, \xi_3) \\ x_2 &= x_2(\xi_1, \xi_2, \xi_3) \\ x_3 &= x_3(\xi_1, \xi_2, \xi_3) \end{aligned} \tag{1}$$

As an example, we consider cylindrical coordinates $\xi_1 = r$, $\xi_2 = \vartheta$, $\xi_3 = z$ (Fig. 1):

$$x_1 = r \cos \vartheta, \; x_2 = r \sin \vartheta \quad x_3 = z$$

The rectangular base (unit vectors) that is associated with the new system is defined as

$$\boldsymbol{e}_{\xi_i} = \frac{\partial \boldsymbol{x}}{\partial \xi_i} \Big/ \left| \frac{\partial \boldsymbol{x}}{\partial \xi_i} \right| = h_i \frac{\partial \boldsymbol{x}}{\partial \xi_i}, \quad i = 1,2,3 \tag{2}$$

With $d\boldsymbol{x} = \sum_{i=1,2,3} \frac{\partial \boldsymbol{x}}{\partial \xi_i} d\xi_i$ (3)

one finds for the line element ds, using

$$\begin{aligned} d\boldsymbol{x} &= h_1 \, d\xi_1 \, \boldsymbol{e}_{\xi_1} + h_2 \, d\xi_2 \, \boldsymbol{e}_{\xi_2} + h_3 \, d\xi_3 \, \boldsymbol{e}_{\xi_3} \\ ds^2 &= (d\boldsymbol{x}, d\boldsymbol{x}) = \sum_{i=1,2,3} h_i^2 \, d\xi_i^2 \end{aligned} \tag{4}$$

whereby $h_i^2 = \sum_{j=1,2,3} \left(\frac{\partial x_j}{\partial \xi_i} \right)^2$. (5)

In case of cylindrical coordinates, one obtains accordingly

$$\begin{aligned} d\boldsymbol{x} &= dr \, \boldsymbol{e}_r + r \, d\vartheta \, \boldsymbol{e}_\vartheta + dz \, \boldsymbol{e}_z \\ (h_1 &= h_3 = 1, \quad h_2 = r) \end{aligned} \tag{6}$$

In continuum mechanics, field quantities are used, for example, a flow velocity field of a fluid, $\boldsymbol{v}(\boldsymbol{x})$ (at each spatial location $\boldsymbol{x}$ the fluid has a particular flow velocity $\boldsymbol{v}$) or a density $\rho(\boldsymbol{x})$. Thereby, methods involving differential calculus, in particular the operations denoted as “gradient,” “divergence,” “rota-

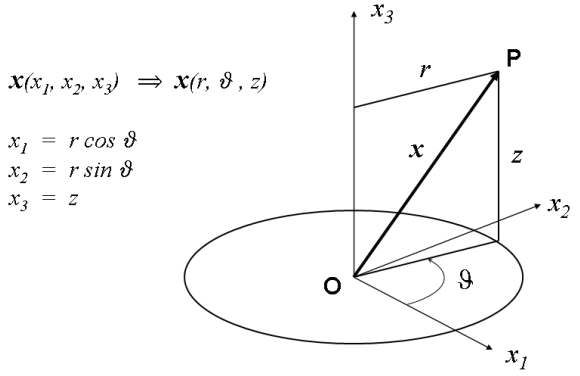


FIGURE 1. Definition of cylindrical coordinates (r, ϑ, z). $\boldsymbol{x}$: Arbitrary vector in space, origin O.

tion," and "Laplace" play a central role. They can be applied to vector and tensor fields, the operations gradient, and Laplace also to scalar fields.

The components of the differentiation operator "gradient," ∇ (a vector expression, often written as *grad*) is defined as, using Equation 4,

$$(\nabla ...) = \left(\frac{1}{h_1}\frac{\partial ...}{\partial \xi_1}, \frac{1}{h_2}\frac{\partial ...}{\partial \xi_2}, \frac{1}{h_3}\frac{\partial ...}{\partial \xi_3} \right) \tag{7}$$

The three dots indicate an element (scalar quantity, components of a vector or tensor field) that the operator acts upon. In Cartesian and cylindrical coordinates, respectively,

$$\begin{aligned} (\nabla ...) &= \left(\frac{\partial ...}{\partial x_1}, \frac{\partial ...}{\partial x_2}, \frac{\partial ...}{\partial x_3} \right) \\ (\nabla ...) &= \left(\frac{\partial ...}{\partial r}, \frac{1}{r}\frac{\partial ...}{\partial \vartheta}, \frac{\partial ...}{\partial z} \right) \end{aligned} \tag{8}$$

In the case of vectors or tensors, expressed in curvilinear coordinates, it has to be noted that the derivative also has to be applied on the base vectors in addition to the components, because their direction, unlike the one of Cartesian base vectors, may change as a function of the location in space. For example, for a vector field $\boldsymbol{v}(\boldsymbol{x}) = \boldsymbol{v}_r\, \boldsymbol{e}_r + \boldsymbol{v}_\vartheta\, \boldsymbol{e}_\vartheta + \boldsymbol{v}_z\, \boldsymbol{e}_z$ (in cylindrical coordinates, Fig. 2) this procedure yields for the base vectors:

$$\begin{aligned} &\frac{\partial \boldsymbol{e}_r}{\partial r} = 0, \quad \frac{\partial \boldsymbol{e}_r}{\partial \vartheta} = \boldsymbol{e}_\vartheta, \quad \frac{\partial \boldsymbol{e}_r}{\partial z} = 0 \\ &\frac{\partial \boldsymbol{e}_\vartheta}{\partial r} = 0, \quad \frac{\partial \boldsymbol{e}_\vartheta}{\partial \vartheta} = -\boldsymbol{e}_r, \quad \frac{\partial \boldsymbol{e}_\vartheta}{\partial z} = 0 \\ &\frac{\partial \boldsymbol{e}_z}{\partial z} = \frac{\partial \boldsymbol{e}_\vartheta}{\partial z} = \frac{\partial \boldsymbol{e}_z}{\partial z} = 0 \end{aligned} \tag{9}$$

The operation "gradient" increases the degree of a field; that is, when applied to a scalar field, a vector field is obtained, when applied to a vector field, a second-degree tensor field results, etc. In contrast, the divergence $(\nabla,...)$ (often written as *div* …) represents a scalar product and reduces the degree of a field, it can therefore only be applied on vector or tensor fields. In Cartesian coordinates, e.g., the divergence of a vector field, $\boldsymbol{v}(\boldsymbol{x})$ is a scalar,

$$(\nabla, \boldsymbol{v}) = \frac{\partial v_x}{\partial x_1} + \frac{\partial v_y}{\partial x_2} + \frac{\partial v_z}{\partial x_3} \tag{10}$$

The rotation, $(\nabla \times \boldsymbol{v})$ (often written as *rot* $\boldsymbol{v}$, again, $\boldsymbol{v}$ denotes an arbitrary vector field; it can also be applied

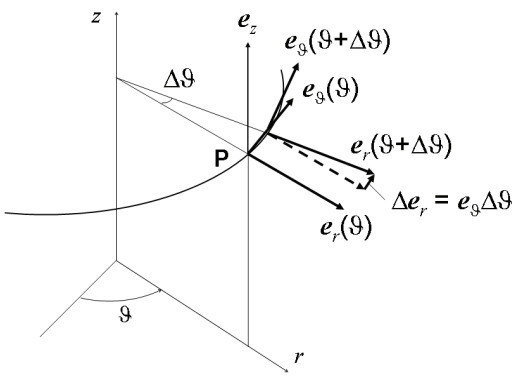


FIGURE 2. Derivatives of base vectors (cylindrical coordinates). When the point P is displaced by Δr in radial direction, the base vectors ($\boldsymbol{e}_r$, $\boldsymbol{e}_\vartheta$, $\boldsymbol{e}_z$) undergo a pure translation; accordingly, all derivatives $\left(\frac{\partial \boldsymbol{e}_i}{\partial r}, i = r, \vartheta, z \right)$ are zero. The same holds for $z \rightarrow z + \Delta z$. Under a variation $\vartheta \rightarrow \vartheta + \Delta\vartheta$, however, the vectors $\boldsymbol{e}_r$ and $\boldsymbol{e}_\vartheta$ change their direction whereby $\Delta \boldsymbol{e}_r = \boldsymbol{e}_\vartheta \Delta\vartheta$ (figure) and $\Delta \boldsymbol{e}_\vartheta = -\boldsymbol{e}_r \Delta\vartheta$.

with tensors) is a vector product. It does not change the degree of a field. In Cartesian coordinates,

$$(\nabla \times \boldsymbol{v}) = \begin{pmatrix} \frac{\partial v_{x_2}}{\partial x_3} - \frac{\partial v_{x_3}}{\partial x_2} \\ \frac{\partial v_{x_3}}{\partial x_1} - \frac{\partial v_{x_1}}{\partial x_3} \\ \frac{\partial v_{x_1}}{\partial x_2} - \frac{\partial v_{x_2}}{\partial x_1} \end{pmatrix} \tag{11}$$

Finally, the "Laplace" operator, $\Delta = (\nabla, \nabla) = \nabla^2$ is, as a scalar product, itself a scalar and does not change the order of a field. In Cartesian and cylindrical coordinates, respectively, the Laplace operator reads

$$\begin{aligned} \nabla^2 ... &= \frac{\partial^2 ...}{\partial x_1^2} + \frac{\partial^2 ...}{\partial x_2^2} + \frac{\partial^2 ...}{\partial x_3^2} \\ \nabla^2 ... &= \frac{1}{r}\frac{\partial}{\partial r}\left(r \frac{\partial ...}{\partial r} \right) + \frac{1}{r^2}\frac{\partial^2 ...}{\partial \vartheta^2} + \frac{\partial^2 ...}{\partial z^2} \end{aligned} \tag{12}$$

Analogous formulas can be derived for other rectangular coordinate systems. A comprehensive treatment is presented in Moon & Spencer[1]; mathematical physics and differential equations are also discussed in Courant & Hilbert[2], as well as in Schaum's Outline.[3]

II. METHODS BASED ON RIGID BODY APPROXIMATIONS

A. Kinematics

An unrestrained point mass in space has three degrees of freedom. These can be represented in the form of the position vector, $\boldsymbol{r}$, in an inertial reference system. If a Cartesian coordinate system with orthogonal unit vectors $\boldsymbol{e}_{x_1}$, $\boldsymbol{e}_{x_2}$, $\boldsymbol{e}_{x_3}$ is chosen, $\boldsymbol{r} = x_1 \boldsymbol{e}_{x_1} + x_2 \boldsymbol{e}_{x_2} + x_3 \boldsymbol{e}_{x_3}$, while the representation in components is $(\boldsymbol{r}) = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$. As the point mass moves in time (t), kinematic quantities derived thereof are the velocity $\boldsymbol{v}(t) = \frac{d\boldsymbol{r}(t)}{dt}$ and the acceleration $\boldsymbol{a}(t) = \frac{d^2\boldsymbol{r}(t)}{dt^2}$.

An unrestrained rigid body has six degrees of freedom. Accordingly, these are parameterized by six quantities, denoted as generalized position coordinates. There are various choices of parameterization, the particular parameterization of use is determined according to the problem to be solved. For example, three linear Cartesian coordinates x_1, x_2, x_3 denoting the location of the center of mass $\boldsymbol{r} = x_1 \boldsymbol{e}_{x_1} + x_2 \boldsymbol{e}_{x_2} + x_3 \boldsymbol{e}_{x_3}$ in a reference system with orthogonal unit vectors as above, and three angles (various options will be specified later) can be chosen. Thus, the body is characterized completely by its position and orientation in space with respect to an inertial reference system. The shape or outer contour of the body has to be considered when collision problems are to be solved, whereas the interior structure can generally not be taken into account. If the number of degrees of freedom is reduced due to external constraints such as, for example, are given by a joint, the number of generalized coordinates is reduced accordingly.

As we shall see later, it is often advantageous to introduce other, in particular non-inertial reference systems, for example, a body-fixed system. However, such a system changes its location and spatial orientation with time. For a complete description of a motion it is therefore necessary to relate the system in use to inertial reference. Accordingly, we have to establish transformations between coordinate systems, which may depend on time.

Two orthogonal coordinate systems can always be made to coincide at a fixed point in time by a translation and a rigid rotation in that, first, the two origins are brought together by a translation; second, a sequence of rotations is made such that the axes (or direction of base vectors) coincide. While the translation is given by a spatial vector, rotations are described by linear mappings that can be represented in matrix form. In order to investigate the properties of such matrices, we make use of the fact that a given vector, $\boldsymbol{c}$, has the same length $|\boldsymbol{c}| = \sqrt{(\boldsymbol{c}, \boldsymbol{c})}$ in any coordinate system (inertial or arbitrarily moving in space).

The components of the vector $\boldsymbol{c}$, written after rotation (Fig. 3), $\begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix}$ and $\begin{pmatrix} c_1' \\ c_2' \\ c_3' \end{pmatrix}$, respectively, are related by a linear mapping:

$$\begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} d_{11} & d_{12} & d_{13} \\ d_{21} & d_{22} & d_{23} \\ d_{31} & d_{32} & d_{33} \end{pmatrix} \cdot \begin{pmatrix} c_1' \\ c_2' \\ c_3' \end{pmatrix} \tag{13}$$

or $\boldsymbol{c} = \boldsymbol{D} \cdot \boldsymbol{c}'$. A first property of $\boldsymbol{D}$ is obtained from the condition of invariant length, that is,

$$(\boldsymbol{c}, \boldsymbol{c}) = ((\boldsymbol{D} \cdot \boldsymbol{c}'), (\boldsymbol{D} \cdot \boldsymbol{c}')) \tag{14}$$

from which follows $(\boldsymbol{D})^{\mathsf{T}} (\boldsymbol{D}) = (\boldsymbol{I})$ upon application of the rules for matrix manipulation (Equation 14a; $\boldsymbol{I}$ denotes the identity tensor), as well as $det(\boldsymbol{D}) = 1$ (Equation 14b). Accordingly, the matrix is orthogonal and its determinant is equal to 1.

If $\boldsymbol{D}$ relates, e.g., the components of a vector in a body fixed to those in an inertial reference system, it will in general depend on time. It can therefore be differentiated with respect to time, and one finds from Equation 14a

$$\frac{d}{dt}\left[(\boldsymbol{D})^{\mathsf{T}} (\boldsymbol{D})\right] = (\boldsymbol{D})^{\mathsf{T}} (\dot{\boldsymbol{D}}) + (\dot{\boldsymbol{D}})^{\mathsf{T}} (\boldsymbol{D}) = (\boldsymbol{D})^{\mathsf{T}} (\dot{\boldsymbol{D}}) + \left[(\boldsymbol{D})^{\mathsf{T}} (\dot{\boldsymbol{D}})\right]^{\mathsf{T}} = 0 \tag{15}$$

Upon defining $(\boldsymbol{D})^{\mathsf{T}} (\dot{\boldsymbol{D}}) = (\boldsymbol{\Omega})$, Equation 15 yields $\boldsymbol{\Omega} = -\boldsymbol{\Omega}^{-1}$ which implies that $\boldsymbol{\Omega}$ can be written in components as

$$(\boldsymbol{\Omega}) = \begin{pmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{pmatrix} \tag{16}$$

(The choice of the signs and of the indices will be justified later.) From the definition of one finds furthermore the important relation

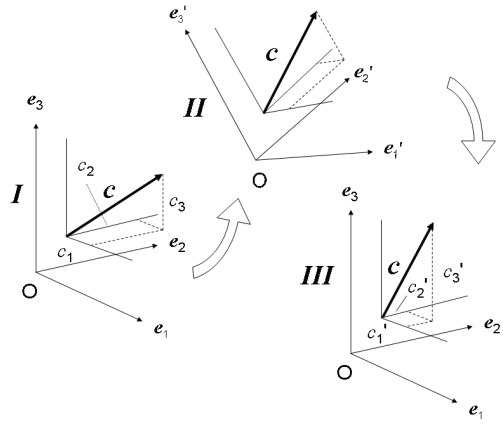


FIGURE 3. Body-fixed reference. The vector $\boldsymbol{c}$ is assumed to be part of a rigid body. Starting from an original configuration (*I*), given in an inertial reference system ($\boldsymbol{e}_1$, $\boldsymbol{e}_2$,$\boldsymbol{e}_3$), the body along with the vector c is moves in space (arrow). In a correspondingly moving body-fixed reference system ($\boldsymbol{e}_1'$, $\boldsymbol{e}_2'$, $\boldsymbol{e}_3'$, *II*) the configuration (as seen from the body-fixed system) remains constant. By execution of a rigid rotation (inverse arrow), the components of the vector can be transformed into inertial reference (*III*).

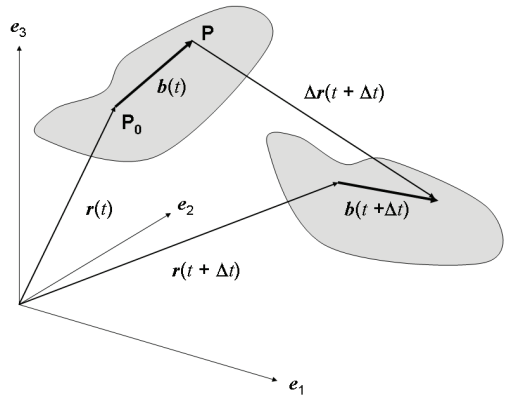


FIGURE 4. Rigid body motion (see text).

$$\dot{\boldsymbol{D}} \;=\; \boldsymbol{D}\cdot\boldsymbol{\Omega} \tag{17}$$

We now look at a fixed point, **P**, on the moving body at time t and $t+\Delta t$, respectively. The body fixed vector, $\mathbf{PP_0}$ is denoted by $\boldsymbol{b}$. From Figure 4 it is seen that the relation holds

$$\boldsymbol{r}(t) + \boldsymbol{b}(t) + \Delta\boldsymbol{r}(t+\Delta t) = \boldsymbol{r}(t+\Delta t) + b(t+\Delta t)$$

$$\text{or}\quad \boldsymbol{r}(t)+\boldsymbol{b}(t)+\frac{d\left[\Delta\boldsymbol{r}(t)\right]}{dt}\Delta t \;=\; \boldsymbol{r}(t)+\frac{d\left[\boldsymbol{r}(t)\right]}{dt}\Delta t \;+\boldsymbol{b}(t)+\frac{d\left[\boldsymbol{b}(t)\right]}{dt}\Delta t$$

for small Δt, implying that

$$\frac{d\left[\Delta\boldsymbol{r}\right]}{dt} \;=\; \boldsymbol{v}\left(\mathbf{P}\right) \;=\; \frac{d\,\boldsymbol{r}}{dt}+\frac{d\,\boldsymbol{b}}{dt} \;=\; \boldsymbol{v}\left(\mathbf{P_0}\right)+\frac{d\boldsymbol{b}}{dt}$$

However, using $\dot{\boldsymbol{b}} = \dot{\boldsymbol{D}}\cdot\boldsymbol{b} = \boldsymbol{D}\cdot\boldsymbol{\Omega}\cdot\boldsymbol{b}$ and observing that for $\Delta t \to 0$ $\boldsymbol{D} \to \boldsymbol{I}$, we arrive at $\boldsymbol{v}(\mathbf{P}) = \boldsymbol{v}(\mathbf{P_0}) + \boldsymbol{\Omega}\cdot\boldsymbol{b}$ or, using component notation,

$$(\boldsymbol{v}(\mathbf{P})) = (\boldsymbol{v}(\mathbf{P_0})) + (\boldsymbol{\omega}\times\boldsymbol{b}) \tag{18}$$

with the angular velocity vector $\boldsymbol{\omega}$ with components $(\boldsymbol{\omega}) = \begin{pmatrix}\omega_1\\ \omega_2\\ \omega_3\end{pmatrix}$. The definition of the components of $\boldsymbol{\Omega}$ has previously made such that the angular velocity vector appears in the well known form given above.

We turn our attention now to the parameterization of the orthogonal matrix $\begin{pmatrix}d_{11} & d_{12} & d_{13}\\ d_{21} & d_{22} & d_{23}\\ d_{31} & d_{32} & d_{33}\end{pmatrix}$. Because the group of spatial rotations is noncommutative, care has to be taken in the choice of an appropriate form, and in fact various formulations have been used in the past depending on the problem at hand.

The oldest form, useful to describe gyroscopes, was introduced by Leonard Euler: The three Euler angles (Fig. 5) (φ, θ, ψ) characterize a rotation around a body axis (self rotation, $\dot{\psi}$, if the body has the largest principal axis at rest in the x_3 direction), a precession ($\dot{\varphi}$), and a nutation ($\dot{\theta}$). It is thereby implicitly assumed that the rotations are related to a fixed reference on the body. The rotations are performed in the sequence $\varphi \to \theta \to \psi$. Accordingly, the first rotations is expressed as

$$\left(\boldsymbol{D}_{\varphi}\right) \;=\; \begin{pmatrix} cos\,\varphi & sin\,\varphi & 0\\ -sin\,\varphi & cos\,\varphi & 0\\ 0 & 0 & 1\end{pmatrix} \tag{19}$$

The matrices associated with the rotations around θ and ψ can be found by interchanging the rows and columns and replacing the angles. Upon multiplication, one finds for the total rotation, $\boldsymbol{D}$, in components,

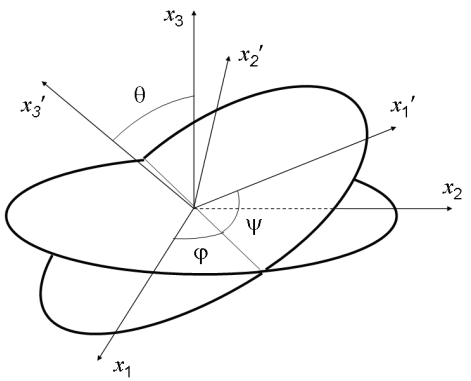


FIGURE 5. Definition of the Euler angles φ, ψ, θ.

$$(\boldsymbol{D}) = \begin{pmatrix} \cos\psi\cos\varphi - \cos\theta\sin\varphi\sin\psi & \cos\psi\sin\varphi + \cos\theta\cos\varphi\sin\psi & \sin\psi\sin\theta \\ -\sin\psi\cos\varphi - \cos\theta\sin\varphi\cos\psi & -\sin\psi\sin\varphi + \cos\theta\cos\varphi\cos\psi & \cos\psi\sin\theta \\ \sin\theta\sin\varphi & -\sin\theta\cos\varphi & \cos\theta \end{pmatrix} \quad (20)$$

A problem derives from the fact that the angles φ and ψ are not defined for $\theta = 0$. Therefore, the Euler angles are not particularly useful for the description of spatial rigid body motions, particularly if numerical applications are made.

The sometimes used expressions "*yaw, pitch, roll*" are adapted from nautical language and denote rotations around the three orthogonal axes of a fixed reference system. Again, the sequence of rotations has to be specified, and ($\boldsymbol{D}$) can be found by multiplication of three successive rotations analogous to Equation 19. This procedure does not have the "not-defined" problem of the Euler angles and is well suited for numerical applications; however, a 2π periodicity remains (as for all rotations).

If the origins of the two coordinate systems in question coincide (which can be achieved by a translation), one can always find a spatial axis $\boldsymbol{\mu}, (|\boldsymbol{\mu}| = 1)$ such that the two coordinate systems coincide also with their axes after performing a (finite) rotation with angle ϑ around $\boldsymbol{\mu}$. From Figure 6 we find that

$$(\boldsymbol{D}) = (\boldsymbol{I})\cos\vartheta - (\boldsymbol{I}\times\boldsymbol{\mu})\sin\vartheta + (\boldsymbol{\mu}\otimes\boldsymbol{\mu})(1-\cos\vartheta) \quad (21)$$

Therefore, the symbol ⊗ denotes a dyadic product and $(\boldsymbol{I}\times\boldsymbol{\mu}) = \begin{pmatrix} 0 & -\mu_3 & \mu_2 \\ \mu_3 & 0 & -\mu_1 \\ -\mu_2 & \mu_1 & 0 \end{pmatrix}$ in Cartesian coordinates.

The four quantities ($\boldsymbol{\mu}$), ϑ are not independent because of $|\boldsymbol{\mu}| = 1$. With the aid of the expression $\varepsilon_{ijk} = \frac{1}{2}(i-j)(j-k)(k-i),\ i,j,k = 1,2,3,$ as well as

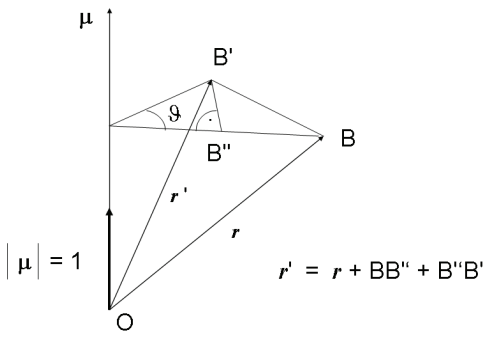


FIGURE 6. Representation of motion by finite rotation: The vector r' is obtained from r by a finite rotation (angle ϑ) around the axis which is given by the unit vector μ. During the entire motion to be described, $\vartheta(t)$ and $\mu(t)$ always refer to the same original configuration.

the Kronecker symbol δ_{ij}, the components of the matrix associated with $\boldsymbol{D}$ are obtained as

$$d_{ij} = \delta_{ij}\cos\vartheta - \sum_k \varepsilon_{ijk}\mu_k \sin\vartheta + \mu_i\mu_j(1-\cos\vartheta) \quad (22)$$

Note: This representation should not be confounded with the concept of the momentary or helical axis, which describes the instantaneous rotation of a rigid body. For further information with respect to this representation, which is sometimes used especially to describe primarily planar, (e.g., the jaw motion), the reader is referred to the literature.[4,5]

A further useful relation is based on direction cosines. Two coordinate systems with base vectors $\boldsymbol{e}_i$ and $\boldsymbol{e}'_i$, respectively, are related by the cosines $c_{ij} = (\boldsymbol{e}_i, \boldsymbol{e}'_j)$ in the form

$$(\boldsymbol{D}) = \begin{pmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{pmatrix} \quad (23)$$

This relation is found by decomposition of the base vectors of one system with respect to the other. (It should be noted that the matrix associated with the transformation of the base vectors and the one with the transformation of the components of a vector according to Equation 13 are inverse. This is particularly important when non-orthogonal systems are used and covariant and contravariant components have to be distinguished.)

For completeness, two further parametrizations are mentioned. A representation using quaternions for rotations is obtained if the definitions

$$\varepsilon = \boldsymbol{\mu} \sin\left(\frac{\vartheta}{2}\right) \text{ and } \varepsilon_4 = \cos\left(\frac{\vartheta}{2}\right) \tag{24}$$

are introduced. Because $\sum_{i=1,\ldots,4} \varepsilon_i^2 = 1$, the 4 quantities (three components of ε and ε_4) have in fact the properties of quaternions. If this representation is used in computational applications, the quaternion property has to be restored after each integration step. However, this disadvantage is offset by the fact that this involves a fitting procedure that adds to the stability of the integration.

Directly related with the quaternions are the three Rodrigues parameters

$$\rho_i = \frac{\varepsilon_i}{\varepsilon_4} \quad i = 1, 2, 3 \tag{25}$$

Because these parameters can become infinite, they are usually not used in numerical procedures.

Two of the representations outlined above (rotation by angle ϑ around axis $\boldsymbol{\mu}$ and direction cosines) are particularly well-suited for numerical applications that involve the integration of linear and angular accelerations. Accordingly, the time derivative of $\boldsymbol{D}$ or an expression for $\boldsymbol{\omega}$, respectively, are needed in terms of the parameters. While in the case of the direction cosine parameterization, Equation 17 can be used directly, Equations 17 and 21 yield, after some calculation,

$$\begin{aligned} \frac{d\vartheta}{dt} &= (\boldsymbol{\omega}, \boldsymbol{\mu}) \\ \frac{d(\boldsymbol{\mu})}{dt} &= \frac{1}{2}(\boldsymbol{\omega} \times \boldsymbol{\mu}) - \frac{\sin\vartheta}{2(1-\cos\vartheta)}((\boldsymbol{\omega} \times \boldsymbol{\mu}) \times \boldsymbol{\mu}) \end{aligned} \tag{26}$$

Once $\boldsymbol{\omega}$ is determined from the equations of motion (see below), the direction matrices can be updated along with the position vectors.

The formal mathematical description in more detail of systems of rigid bodies in motion can be found in Kane[6,7] and Wittenburg,[8] which also include the dynamics (Chapter C).

B. Measurement Methods

The measurement of human body motions is either based on optical methods (position measurement of body surface markers, mostly utilizing infrared, video, or film recording along with associated image analysis) or on the application of acceleration sensors (Fig. 7). While velocities and accelerations from position measurements are obtained by differentiation, accelerations have to be integrated in order to determine velocities and positions. In the case of differentiation, noise amplification is a well-known problem, and appropriate filters have to be applied in order to ascertain extraction of the significant signal contents. Linear accelerometers, which measure accelerations in one direction, can be made small and lightweighted; the measurement of angular or rotational accelerations is in contrast more critical and has for a long time posed technical problems, in particular in trauma research where high levels are reached. Today, however, angular accelerometers for various measurement ranges and purposes are available.[9-11] Integration of accelerations is not straightforward because the group of rotations is not commutative and, furthermore, drift problems occur.

Optical methods for position measurements are often based on close-range photogrammetry[12-16] and are widely used for clinical gait analysis as well as for computer-assisted surgery navigation.[17-19] In a close-range photogrammetry setting, various cameras are installed to monitor a laboratory, surgical theatre, or generally an object space. A point **A** in the object space with coordinates (X_A, Y_A, Z_A) is thereby imaged (Fig. 8) on the 2D camera space of camera i, image coordinates (x^i_A, y^i_A). The photogrammetric procedure consists of the determination of the

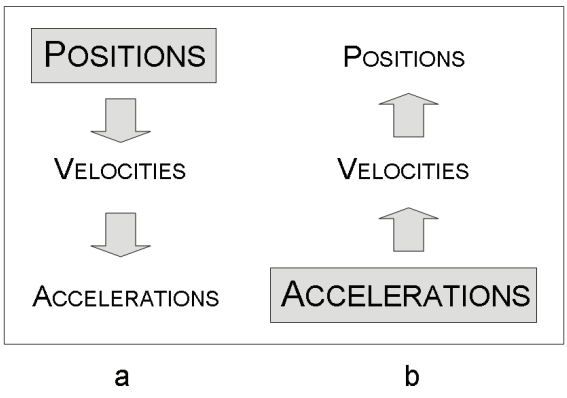


FIGURE 7. Experimental determination of positions, velocities, and accelerations. When position measurements are made (a), velocities and accelerations are obtained from differentiation, if accelerations are measured (b), integrations are necessary.

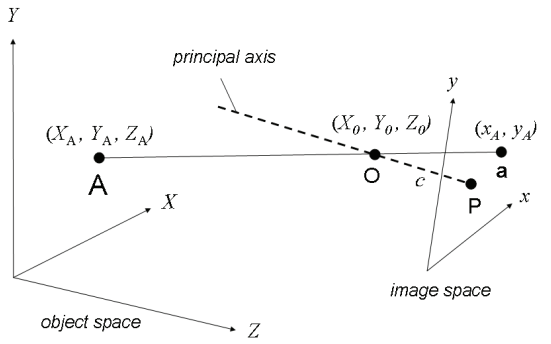


FIGURE 8. Close range photogrammetry for motion analysis. The 3D object space (representative, arbitrarily chosen point A, object space coordinates X_A, Y_A, Z_A) is imaged onto the 2D image space, point a with coordinates x_A , y_A. The principal axis corresponds ideally to the optical axis and is perpendicular to the image space. The principal point P denotes the section of the principal axis through the image space.
O: Perspective center. This is a camera-fixed point, given by the imaging properties of the lens. Its coordinates are X_0, Y_0, Z_0 ; these are given in object space coordinates and relate object and image space. The respective coordinate systems are connected to one another by a rigid rotation (matrix r_{kl}) and a scaling factor, μ (see text).
c: Camera constant (distance between the principal point P and the perspective center). In a typical close-range photogrammetry setting, this quantity is close to the image-sided focal distance of the lens. More than one camera (with non-parallel principal axes) are needed to facilitate recovering 3D object space coordinates from the planar image coordinates.

spatial location of **A** from the planar camera views, that is, from the coordinates (x^i_A, y^i_A). The general procedure is as follows: for each camera *i*, the two coordinate systems (object space and image space of the camera, respectively) are related to one another by a rotation (matrix r^i_{kl}). $\mathbf{O}^i$ denotes furthermore the perspective center with coordinates $\left(X^i_0, Y^i_0, Z^i_0\right)$. Note that these camera-specific coordinates refer to the object space. With the scaling factor, μ^i, the relation holds (similarity of two triangles after rotation)

$$\begin{pmatrix} x^i_A \\ y^i_A \\ -c^i \end{pmatrix} = \frac{1}{\mu^i} \begin{pmatrix} r^i_{11} & r^i_{12} & r^i_{13} \\ r^i_{21} & r^i_{22} & r^i_{23} \\ r^i_{31} & r^i_{32} & r^i_{33} \end{pmatrix} \cdot \begin{pmatrix} X_A - X^i_0 \\ Y_A - Y^i_0 \\ Z_A - Z^i_0 \end{pmatrix} \qquad (27)$$

The quantity c^i is denoted as camera constant. Upon utilizing the third equation to eliminate the scaling factor, one obtains the so-called collinearity equations

$$\begin{aligned} x^i_A &= \frac{-c^i \left(r^i_{11}\left[X^i_0 - X_A\right] + r^i_{12}\left[Y^i_0 - Y_A\right] + r^i_{13}\left[Z^i_0 - Z_A\right] \right)}{r^i_{31}\left(X^i_0 - X_A\right) + r^i_{32}\left(Y^i_0 - Y_A\right) + r^i_{33}\left(Z^i_0 - Z_A\right)} \\ y^i_A &= \frac{-c^i \left(r^i_{21}\left[X^i_0 - X_A\right] + r^i_{22}\left[Y^i_0 - Y_A\right] + r^i_{23}\left[Z^i_0 - Z_A\right] \right)}{r^i_{31}\left(X^i_0 - X_A\right) + r^i_{32}\left(Y^i_0 - Y_A\right) + r^i_{33}\left(Z^i_0 - Z_A\right)} \end{aligned} \qquad (28)$$

The location $\left(X^i_0, Y^i_0, Z^i_0\right)$, the camera constant c^i, along with the matrix r^i_{kl} have to be determined for each camera using measurement grids set up in the laboratory. Theoretically, two cameras are then needed to determine (X_A, Y_A, Z_A) from the planar camera views using Equation 28. In reality, however, the intransparency of the human body requires multiple camera arrangements. Furthermore, the equations have to be corrected for practical inaccuracies; in particular, a careful camera calibration[20] is necessary (the theoretical principal axis may deviate from the actual camera axis, the image space is not accurately perpendicular to it, and, furthermore, lens distortions have to be taken into account).

With respect to procedures based on acceleration measurements, it was mentioned above that the measurement of rotational accelerations posed technical difficulties in the past. It is sometimes attempted to obtain these with the aid of a combination of linear accelerometers, which are small, lightweight, accurate, and have a large frequency response. Because these elements are usually fixed on the human limbs and therefore change their spatial orientation when the body moves, the term "strap-down" measurement is used. Under such conditions, linear and angular accelerations with respect to inertial reference cannot be obtained directly. A problem exists in particular because linear accelerometers always sense gravity and a signal is recorded already when a linear accelerometer is rotated with respect to the direction of gravity. We consider the output of a linear accelerometer that moves in space. From Equation 18, after differentiation with respect to time, we obtain for the acceleration, $\boldsymbol{a}$, measured at an arbitrary fixed point **P** on a rigid body in motion

$$\left(\boldsymbol{a}(\mathbf{P})\right) = \left(\boldsymbol{a}\left(\mathbf{P_0}\right)\right) + \left(\dot{\boldsymbol{\omega}} \times \boldsymbol{b}\right) + \left(\boldsymbol{\omega} \times \boldsymbol{\omega} \times \boldsymbol{b}\right) \qquad (29)$$

where $\mathbf{P_0}$ is a reference point on the same body such that the distance $\mathbf{PP_0} = |\boldsymbol{b}|$ is constant in

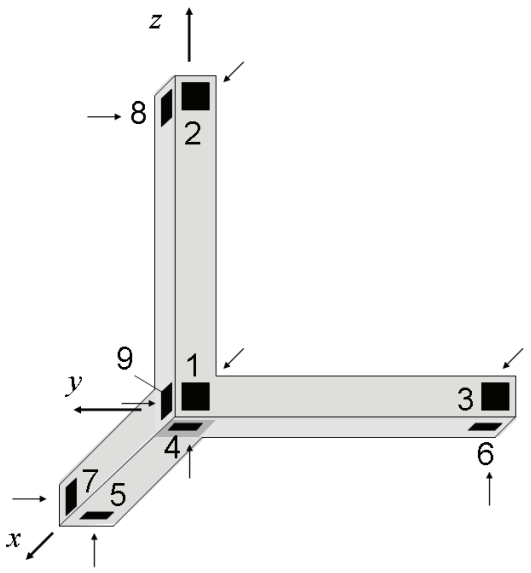


FIGURE 9. Measurement of rotational accelerations. A typical arrangement of nine linear accelerometers that allows determining rotational accelerations. Small arrows indicate positive measurement directions.

time. Theoretically, six independent measurements $\boldsymbol{a}_i(\mathbf{P}_i)$, $i = 1,...6$ would be sufficient to determine $\boldsymbol{a}(\mathbf{P}_0)$ and $\boldsymbol{\omega}$ if the distances $\mathbf{P}_0\mathbf{P}_i$ are known and the points $\mathbf{P}_i$ are chosen appropriately (in particular, collinearity has to be avoided). However, due to the fact that $\boldsymbol{\omega}$ appears in quadratic form in Equation 29, this procedure is associated with intolerable errors in practical applications.

With a minimum of nine linear accelerometers, which are arranged according to the scheme shown in Figure 9, the quadratic terms can be eliminated. To show this, we consider the three measurement components $(\boldsymbol{a})_i = a_i$ in direction x (x_j, y_j, z_j denote the location of the accelerometers), using Equation 29,

$$a_i = a_{x_0} + \dot{\omega}_y z_j - \dot{\omega}_z y_j + \omega_x\left(\omega_y y_j + \omega_z z_j\right), \quad \left(x_j = 0\right) \tag{30}$$

(For each component i the appropriate coordinates x_j, y_j, z_j have to be inserted). The other two directions (y, z) yield similar formulas whereby the sign of the acceleration measurement has to be observed. Upon elimination of the quadratic terms, one arrives at

$$\begin{aligned} 2\dot{\omega}_x &= \frac{a_8 - a_9}{z_8 - z_9} + \frac{a_6 - a_4}{y_6 - y_4} \\ 2\dot{\omega}_y &= \frac{a_2 - a_1}{z_2 - z_1} - \frac{a_5 - a_4}{x_5 - x_4} - \frac{(a_3 - a_1)(y_2 - y_1)}{(y_3 - y_1)(z_2 - z_1)} \\ 2\dot{\omega}_z &= \frac{a_7 - a_9}{x_7 - x_9} + \frac{a_3 - a_1}{y_3 - y_1} \end{aligned} \tag{31}$$

These equations can readily be integrated. However, a drift problem is still persistent. This is not the case if position measurements are made. A drawback nevertheless exists because by the differentiation and necessary filtering process, bandwidth is lost and, accordingly, high-speed recording requiring intensive light is necessary. In contrast, the bandwidth of accelerometers is sufficient for all practical purposes.

C. Dynamics

Most dynamic phenomena of interest in biomechanics are such that their complexity does not allow for an analytic treatment; rather, numerical procedures have to be applied. Accordingly, the following considerations are formulated in view of computer modeling. It is furthermore noted that rigid body equations also apply for deformable bodies with respect to the center of mass.

The modeling and analysis of rigid body motions is based on the linear and angular momentum balance. We consider a system consisting of N rigid segments that may be connected in joints or not. There may be further constraints such as gliding along a plane. For each rigid segment n ($n = 1,..., N$) the following two equations hold:

First, the momentum equation reads

$$m_n \ddot{\boldsymbol{r}}_n = \sum_k \boldsymbol{F}_{nk} + \sum_j \boldsymbol{f}_{nj} \tag{32}$$

with: m_n mass of segment n ; $\boldsymbol{r}_n$ location of associated center of mass; $\boldsymbol{F}_{nk}$ force k acting on segment n ; $\boldsymbol{f}_{nj}$ force of joint j, acting on segment n.
Second, the angular momentum balance (component formulation) in body-fixed reference is (note that this is only correct if the equations are formulated with respect to the center of mass)

$$\begin{aligned} (\boldsymbol{I}_n, \dot{\boldsymbol{\omega}}_n) = &-(\boldsymbol{\omega}_n \times (\boldsymbol{I}_n, \boldsymbol{\omega}_n)) + \\ &\left(\boldsymbol{D}_n, \left[\sum_k \left(\boldsymbol{D}_n^{-1}, (\boldsymbol{\rho}_{nk} \times \boldsymbol{F}_{nk})\right) + \sum_i \left(\boldsymbol{D}_n^{-1}, (\boldsymbol{\kappa}_{ni} \times \boldsymbol{f}_{ni})\right) + \sum_j \left(\boldsymbol{D}_n^{-1}, \boldsymbol{t}_{nj}\right) + \sum_m (\boldsymbol{\tau}_{nm})\right]\right) \end{aligned} \tag{33}$$

with: $\boldsymbol{I}_n$ inertia tensor of segment n in body-fixed reference (with respect to the center of mass); $\boldsymbol{\omega}_n$ angular velocity of segment n ; $\boldsymbol{D}_n^{-1}$ transformation from body-fixed into inertial system; ρ_{nk} point of application of external force $\boldsymbol{F}_{nk}$; $\boldsymbol{\kappa}_{ni}$ point of application of joint force $\boldsymbol{f}_{ni}$; $\boldsymbol{t}_{nj}$ torque j, acting on segment n due to joint constraints; $\boldsymbol{\tau}_{nm}$ external torque m, acting on segment n.

The term on the left side of Equation 33 and the first term on the right side originate from

$$\frac{d}{dt}\left(\boldsymbol{D}_n^{-1},\left(\boldsymbol{I}_n,\boldsymbol{\omega}_n\right)\right) \tag{34}$$

where the relation in Equation 17 has been applied. This procedure is useful because the inertia tensor $\boldsymbol{I}_n$ is constant in a body-fixed reference system. To include external torques, the equation is written in inertial reference, but it is finally multiplied with $\boldsymbol{D}_n$ in order to transform it into body-fixed reference.

These two basic equations are valid for every system of rigid bodies. In practical applications, there are in addition various types of constraints due to joints (ball and socket, hinge, etc.) and external conditions (e.g., sliding along a plane) that are special for the system under consideration. All of these constraints are usually assumed to appear as functions of the form

$$f\left(\boldsymbol{D}_n,\boldsymbol{r}_n,\dot{\boldsymbol{r}}_n,angles,\boldsymbol{\omega}_n,external\ parameters\right) = 0 \tag{35}$$

with first-order time derivatives at most (otherwise, the mathematical character of the equations might change). In the case of viscoelastic damping, a hereditary integral may also be involved. These equations are added to the basic equations of motion and the entire set can directly be utilized for numerical integration. The constraints reduce the degree of freedom of the system and the first step during the integration procedure of Equations 32, 33, and 35 consists therefore of eliminating "superfluous" variables. Although this complicates the method, an important advantage is due to the flexibility that is maintained in that changes of the model, of the connectivity of a rigid body system, of joint properties, or of external constraint can readily be implemented. Likewise, quantities in inertial or body-fixed reference are easily calculated.

An alternative approach for the formulation of the basic equations is based on the Lagrangian formulation. The number of degrees of freedom M of the system consisting of N rigid segments has first to be determined taking into account all of the constraints, and every degree of freedom has to be associated with a generalized position coordinate, q_k ($k = 1,...,M$). For every degree of freedom the equation of motion (Lagrange equation) then reads

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_k}\right) - \frac{\partial T}{\partial q_k} = \frac{\partial V}{\partial q_k} + Q_k \quad (k = 1,...,M) \tag{36}$$

T denotes the kinetic energy, $2T = \sum_{n=1}^{N} m_n\left(\boldsymbol{v}_n,\boldsymbol{v}_n\right) + \sum_{i=1}^{N}\left(\boldsymbol{\omega}_n,\left(\boldsymbol{I}_n,\boldsymbol{\omega}_n\right)\right)$, and V a potential. The forces Q_k ("generalized forces") include forces that cannot be described by a potential. These have to be derived with the aid of the principle of virtual work. (Note that also internal forces may contribute, e.g., friction.) This approach leads to a minimal set of equations because all the constraints are included in the formulation. The disadvantage derives from the lack of flexibility such that this method is only used if the system to be analyzed is fixed. It is therefore less suited for general purpose programs.

Other procedures have been proposed and can be found in the literature; however, the two methods outlined here are the most important (and are moreover the basis for other formulations). For the numerical integration, an integrator that is adapted for stiff equations is needed (see remark below). After each integration step, the orthogonality of the transformation matrices has usually to be restored by a fitting procedure (in case that Equations 32 and 33 are used). A simple (but not exhaustive) control can be made by following the common center of mass of the system under consideration, which has to execute a linear, constant motion if external forces are absent. Likewise, an energy balance can be performed.

An important aspect that is often underestimated is related to the sensitivity of the calculated results with respect to the initial conditions and system parameters. Small changes in the initial conditions or in the values of the system parameters may sometimes lead to large deviations and variability of the resulting motions, in particular in the case of collisions. (Consider the case of a soccer ball hitting a goalpost: a minute difference in the trajectory of the

ball may decide whether the ball will end in the goal or go astray.) Often, numerous parametric variations are necessary in order to validate the significance and usefulness of a simulation. Systems exhibiting excessive sensitivity to certain parameters or initial conditions may indicate an unsuitably defined model.

Collisions may cause difficulties in that within the framework of rigid bodies, collisions lead to singularities (infinite forces acting over an infinitely short period of time). This may be the reason for stiffness problems mentioned above. There are in essence two ways to deal with this problem. First, the bodies are "softened" such that there are no infinite forces. To this end, a force-penetration model is developed for which the outer contours of the segments are needed. Mutual penetration can readily be calculated from the centers of mass and orientation matrices. Second, each collision configuration is determined exactly (machine precision) by interpolation and the collision is bridged with the aid of conservation laws.

A further procedure consists of including a formulation based on continuum mechanics for parts of special interest. Figure 10 shows the model of a skier that consists of a system of rigid bodies combined with a finite element model (see Chapter G, Numerical Procedures, in III. Continuum Mechanics) representing the knee. This model describes the gross motion and allows determining the forces acting in the system. The effects of the forces in the knee joint, in particular in the various ligaments, are then analyzed by way of a detailed deformable knee model.

Multibody systems are mostly applied in human motion analysis, gait analysis,[21-27] rehabilitation,[28] and trauma research.[5,29] A particular problem is associated with the determination of internal forces (joints) from motion recordings (inverse problem).[30-33]

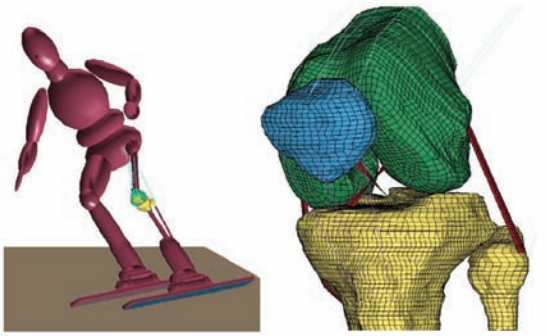


FIGURE 10. Computer model of a skier. For the purpose of motion analysis, the skier is modeled as a system of rigid bodies connected in joints. The behavior of the knee joint under the influence of the forces generated during the motion is analyzed by replacing part of the rigid body system by a continuum-based model (finite element formulation, see Chapter G. Numerical Procedures in III.Continuum Mechanics). (Used with permission from Semadeni R, Schmitt KU. Numerical simulations to assess different rehabilitation strategies after ACL rupture in a skier. J Sport Rehab. 2009;18:427–37.)

D. Molecular Dynamics

Atoms interact with one another primarily through their electron clouds. Their structure is organized in well-defined shells; in the case of incomplete shells, the interaction is particularly strong and molecules are formed (e.g., H_2O, CO_2) through bonds that are denoted as covalent (electron pairs). The carbon atom can also create long chains, which can serve as a backbone for large molecules, in particular biomolecules. Molecular interactions are based on various effects: ionic bonds due to electron exchange, dipole effects, noncovalent bonds, hydrogen bridges, van der Waals forces, hydrophilic-hydrophobic interactions (of particular importance for biomolecules). In the case of metallic materials, there are still other effects, such as excited electronic states and interactions with electromagnetic fields, that have to be taken into account. Finally, spin systems have to be observed.

The electron mass and atomic dimensions are such that the principles of quantum mechanics, in particular the Schrödinger equation, has to be applied for the analysis of atomic systems.[34-36] This equation incorporates the probability or wave function $\Psi(\boldsymbol{r},t)$, which describes the probability that a particle is at a certain location $\boldsymbol{r}$ at time t. The integral over the volume V of interest, as is required for probabilities, follows $\int_V |\Psi|^2 dV = 1$. The Schrödinger equation reads

$$i\hbar \frac{\partial \Psi}{\partial t} = -\frac{\hbar}{2m}\nabla^2\Psi(\boldsymbol{r},t) + V(\boldsymbol{r},t)\Psi(\boldsymbol{r},t) \qquad (37)$$

Where m denotes the mass of the particle whose probability distribution is to be calculated, $\hbar$ is

Planck's constant h divided by 2π, $V(\mathbf{r}, t)$ a potential, and $i = \sqrt{-1}$. To determine the energy states of a system, the time-independent Schrödinger equation is derived by application of the separation ansatz $\Psi(\mathbf{r},t) = \psi(\mathbf{r}) exp\left(-\frac{iEt}{\hbar}\right)$ to Equation 37. This yields the eigenvalue equation for the energy E, given the Hamilton operator $H = -\frac{\hbar}{m}\nabla^2 + V$,

$$H\,\psi(\mathbf{r}) = E\,\psi(\mathbf{r}) \tag{38}$$

Only few analytic solutions of the Schrödinger equation exist. Even small molecules (let alone large biomolecules) represent a multi-body problem that requires enormous computational resources and is numerically difficult to approach. Accordingly, computational chemistry has developed itself into a scientific discipline in its own right with numerous different aims and applications, ranging from pharmaceutics (e.g., drug design) and biology (e.g., protein folding) to physics (e.g., solution of the Schrödinger equation with finite element methods).[37-40] Specialized commercial modeling packages are also available for various applications.

Problems of interest in biomechanics involve mainly the interaction and functional properties of protein molecules, including their behavior under the influence of forces. Many biomolecules can execute their biological (including pathological such as is the case for prions!) function only in a well-defined steric configuration. Thereby, folding of proteins depends essentially on the presence of a large number of water molecules because these molecules are polar (note that the dielectric constant of water is 81). Because proteins may consist of several hundred thousand individual atoms, systems with a huge number of degrees of freedom have to be treated. A further circumstance adds to the complexity of the problem in that typical integration steps in case of dynamic simulations have to be chosen on the order of 10^{-15} to 10^{-14} seconds. This small step size is dictated by the internal dynamics of the molecules, such as vibrations. An enormous number of integration steps therefore has to be made in order to reach useful time spans.

Two major procedures that are used in molecular modeling include molecular dynamics on the one hand and "ab initio" modeling on the other. Standard chemical analysis methods allow determining the composition of a molecule in the form of the chemical formula. By more elaborate methods (x-ray diffraction, nuclear magnetic resonance, Raman spectroscopy) the steric configuration can partly be resolved. Molecular mechanics and dynamics are then used to determine and analyze the properties of the steric configuration in detail. Ab initio calculations, in turn, start from scratch and entire molecules are built theoretically using molecular models.

The Born-Oppenheimer approximation is usually applied as a first step in order to reduce the number of degrees of freedom. Because the mass of an electron is more than three orders of magnitude smaller than the mass of a nuclear particle (proton or neutron), the motion of the nucleus can be assumed to be largely independent from the motion of the electrons. A factorization is therefore possible that separates nuclear and electronic motion. This approximation allows modeling the electrons in a summary fashion as a static cloud and replacing their influence by a potential function that mimics the forces that they cause upon interaction with neighboring atoms.

A further step may be based on the application of Ehrenfest's theorems, according to which the expectation values of the Schrödinger wave function, Ψ, obey the laws of classical mechanics. In case that the de Broglie wavelength $\lambda = \sqrt{\frac{2\pi\hbar^2}{mkT}}$ (k denotes Boltzmann's constant, T the absolute temperature) is small in comparison with the typical dimension of the problem under investigation (e.g., atom-atom distances), molecular dynamics can be adequately described by classical mechanics. Already the mass of a helium atom (~10^{-26} kg) is sufficiently large that this approximation is justified.

Because atomic nuclei are small in comparison with molecular dimensions, "simple" point mass mechanics can be applied for each constituent, i, of a molecule (mass m_i, location $r_i(t)$)

$$m_i \frac{d^2\mathbf{r}_i}{dt^2} = \sum_k \mathbf{F}_{ik} \tag{39}$$

While this equation is rather straightforward, the formulation of the interacting forces ($\sum_k \mathbf{F}_{ik}$, k extends

over all interactions) representing the electrons within the framework of the Born-Oppenheimer approximation is not, because this has to be derived from quantum mechanics or empirical knowledge. A large body of literature can be found on this difficult and crucial subject. In a typical potential, the forces associated with variations in bond length and angles are derived from linear spring potentials. To describe particle interactions (atoms or molecules, index i, j, mutual distance r_{ij}), the Lennard-Jones potential is used

$$V_{ij} = C_1 \left(\left[\frac{C_2}{r_{ij}} \right]^{12} - \left[\frac{C_2}{r_{ij}} \right]^{6} \right) \qquad (40)$$

This potential (Fig. 11) describes the strong repulsion $(\propto \frac{1}{r^{12}})$ that atoms are exposed to when their electron clouds are overlapping and the van der Waals attraction at larger distances $(\propto \frac{1}{r^{6}})$. C_1, C_2 are constants. These have to be fitted or experimentally determined, along with the numerous other constants of the spring potentials mentioned above according to the specific interaction to be modeled. Finally, long-range interactions due to electrostatic forces whose potential exhibit a $1/r$ dependence are added. Accordingly, all particles can in principle interact and many-body interactions can occur. In case of thousands of particles, this leads to an untreatable complexity such that further approximations are necessary.

In ab initio models, the electrons cannot be modeled in a summary fashion, rather, the electrons have to be taken into account individually. To this end, the Hartree-Fock approximation is usually applied, where the electron wave function is decomposed into factors representing each electron while observing the Pauli principle (Slater determinants). The reader is referred to the literature for more detail (A.R. Leach[39]).

As an example for molecular modeling, Figure 12 shows a model of a polypeptide in various stages of penetration through a biological membrane.

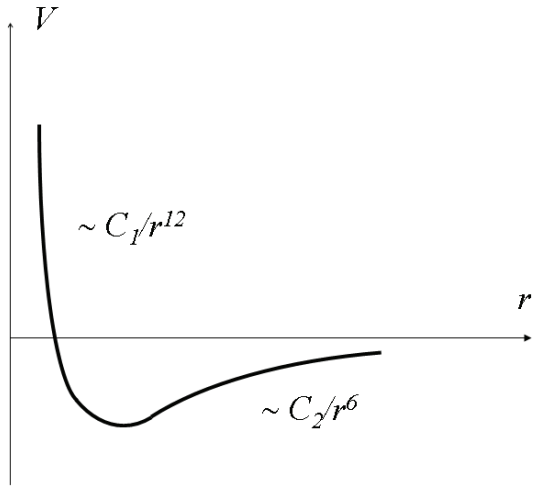


FIGURE 11. Lennard-Jones potential (qualitative).

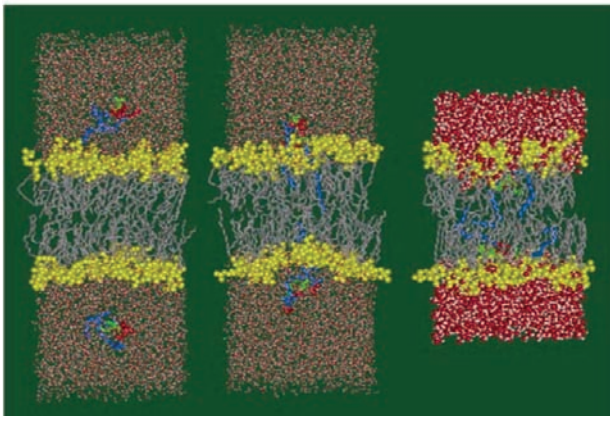


FIGURE 12. Snapshots of the association of a peptide (red-blue-green) in a biological double lipid membrane (yellow-gray), after 0.25 nsec (left), 0.35 nsec (middle), and 0.42 nsec (right). Above and below the membrane are numerous water molecules. (For details, the reader is referred to Gorfe AA, Pellarin R, Caflisch A. Membrane localization and flexibility of a lipidated ras peptide studied by molecular dynamics simulations. J Am Chem Soc. 2004;126:15277–86. Figure reprinted with permission.)

III. METHODS BASED ON CONTINUUM MECHANICS

When the mechanical behavior of tissues, organs, cells, or subcellular structures is to be analyzed in view of their biological structure and function, rigid-body approximations incorporate often insufficient detail. An approach based on continuum mechanics is usually necessary to study tissue deformation patterns; this allows in particular the inclusion of fluid flow simulations, complex transport phenomena, and chemical reactions taking into account electro-

chemistry (the latter are not considered here). Analyses and models aimed toward mechanobiology or mechanotransduction,[40-62] that is, the transduction of mechanical into biological signals, rely especially on such general approaches. This branch of biomechanics has been developing extensively, and accordingly a large body of literature exists covering a wide range of interest, notably relating to osteocytes (bone, skeleton),[63-75] chondrocytes (cartilage),[76-81] endothelial cells,[82-83] fibroblasts,[84] hair cells,[85] (mesynchemal) stem cells,[86,87] animal cells,[88] and more.

Solids and fluids require somewhat different modeling methods, as will be shown in the chapters that follow. Stress-strain characteristics of solid biological tissues are typically nonlinear, anisotropic, and viscoelastic. The nonlinearity manifests itself mainly in cases of large tissue deformations that are often observed in biomechanics, the anisotropy is due to the fibrous character of biological tissues, and the viscoelasticity derives from the internal friction inherent in the fiber-extracellular matrix composition. Fiber rearrangement during deformation may simulate quasi-plastic properties. There are also active elements (muscle fibers and contractile actin elements) whose tone influences the mechanical properties. Further, in tests made under ex vivo conditions, the state of muscle activation has to be taken into account (muscle fibers can be activated chemically, e.g., by Barium compounds). Likewise, embalming of cadavers changes their mechanical behavior. For biological fluids, non-Newtonian characteristics may be important. Complex fluids such as blood contain solid, albeit highly deformable particles (blood cells). In flow conditions in which the conduit diameter and particle size are comparable (e.g., capillaries), flow modeling therefore becomes difficult. A common feature of biological materials is that they are solid-fluid mixtures. Various approaches taking this property into account can be found in the literature (mixture theory).[94,95]

A distinction is often made in biomechanics between "soft" and "hard" tissues. In order to specify this difference more quantitatively, the nonlinear, anisotropic, partly active (muscles) properties of biological tissues have to be characterized by a simplified linear approximation. Under uniaxial loading of a long and thin specimen, a piecewise linear stress-strain relation in the form of Hooke's law can be adapted and a local modulus of elasticity or Young's modulus, E, can be defined. For "soft" tissues, E varies typically between 10 and 10^5 kPa, whereas the values for "hard" tissues are on the order of several gigapascals.

While there are numerous kinds of soft tissues, hard tissue in humans appears essentially in the form of calcified tissue, in particular bone. The calcium is contained in hydroxyapatite crystals ($Ca_5 [PO_4]_3 OH$) that are embedded in a collageneous matrix. Aside from the integrity and mechanical loading capacity of bones, a physiologic calcium balance is eminently important for the overall homeostasis of the human body in that calcium is essential for many physiologic processes, such as the action of muscles, the transmission of nerve signals, and the coagulation of blood. Calcium is by far the most abundant bone mineral material ("calcium reservoir"), and others such as phosphorus are much less concentrated. Therefore, the terms "calcification" and "mineralization" of bone are often used synonymously.

Bone exists in various forms. Cortical bone makes up the shaft (metaphysis) of the long bones as well as the outer layer of other bones, while trabecular or cancellous bone is located mostly in the medullary canal of long bones, particularly in regions close to joints (epiphysis) as well as in the spine and in bones whose primary task is not to support loads (e.g., skull, iliac crest).

The main constituents of soft tissues from a biomechanics point of view are elastin, collagen, and smooth muscle fibers. Elastin, a globular, highly extensible polypeptide (again under the simplified approximation of a piecewise linearized treatment of uniaxial loading) has a Young's modulus of 10^2 to 10^3 KPa, collagen (a stiff three-fold triple helical molecule) a Young's modulus of up to 10^5 KPa, while smooth muscle fibers cover a wide range of stiffness characteristics between elastin and collagen depending on the state of activation. The anatomy and composition of "soft" (non-calcified) organs such as the liver, the brain, muscles, the kidneys are mostly determined by their physiologic function. According to the great variety of physiologic functions, therefore, the composition and microanatomy of soft tissues varies greatly, as does their mechanical behavior under load.

Elasticity derives mostly from the spring-like behavior of materials, in which inter- and intramolecular potentials or potentials of ions that are arranged in a regular crystal or amorphous structure change upon deformation. In the case of large molecules, however, there is also an elasticity component due to the possibility of large configurational changes not involving mechanical potentials. A simple model based on thermodynamic considerations illustrates this phenomenon.

We look at a uniaxial element, length *l*, that is subjected to extension by a force *F* (Fig. 13). According to the second law of thermodynamics, the change in internal energy of the system, *dU* upon elongation by *dl* is

$$dU = T\,dS + F\,dl \tag{41}$$

where *T* denotes the absolute temperature and *S* the entropy of the system. Accordingly, the internal energy depends not only on the action of the mechanical force, *F*, but also on the entropy. For "normal" materials, deformation is not associated with a significant change of entropy. In the case of large molecules, however, this may be different. The simplest model reflecting this finding is a chain molecule that can be compared with a bicycle chain, that is, it consists of numerous links (e.g., monomers) that can freely rotate around their binding sites (Fig. 14). The entropy of such a system is given by the volume that the set of all geometrically possible configurations (end-to-end distance fixed) occupies in the phase space (multidimensional space spanned by all the position coordinates). The farther away the ends of the chain are, the smaller is this volume, implying that the entropy decreases with elongation. This can be found in materials such as highly extensible rubber that are sometimes denoted as "entropy springs."[89-91] In pure form, however, such materials hardly exist, in particular not in biomechanics. Elastin, having a molecular weight of some 65 kD, nevertheless has properties resembling at least partially an entropy spring. Thanks to its globular structure when not loaded, it has an enormous extensibility associated with rubber-like mechanical properties.

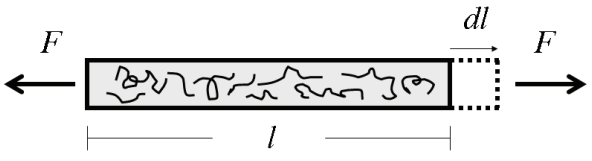


FIGURE 13. Uniaxial loading (Force *F*) and extension ($l \rightarrow l + dl$) of a strip. The composition of the material is not further specified, but homogeneous.

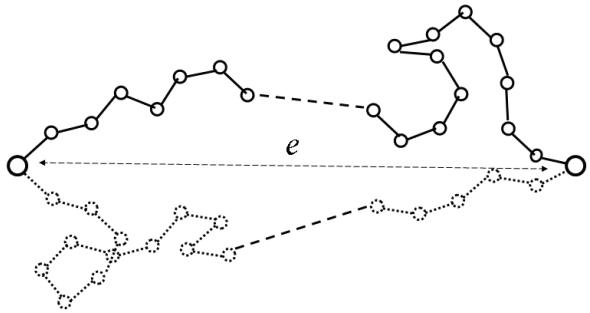


FIGURE 14. "Bicycle"-chain model of an elastomer. Because it is assumed that the links are connected in force-free joints, the chain may adopt any configuration that is geometrically possible between the two ends (fixed end-to-end distance *e*) without change of internal energy.

Forces acting on or in a body and the resulting deformations are related by material-specific properties called constitutive properties. In the case of a continuum, forces and deformations have to be brought into an appropriate form that is useful for calculations.

A. Deformation: Strain

We consider an arbitrary partial or total volume of a continuous, homogeneous solid material in a not further specified, but well-defined initial state that we denote as reference configuration. In the case of homogeneous isotropic fluids without memory effects, particles can be rearranged freely, such that there is no well-defined configuration that could be used as representative initial state. The following derivations with respect to deformations are therefore not directly applicable for fluids.

It is important to note that the reference configuration can be chosen arbitrarily. The body (or part thereof) is in particular not necessarily in a stress-free state; in fact, a solid body does not even need to have a stress-free configuration. We define a Cartesian coordinate system such that every material

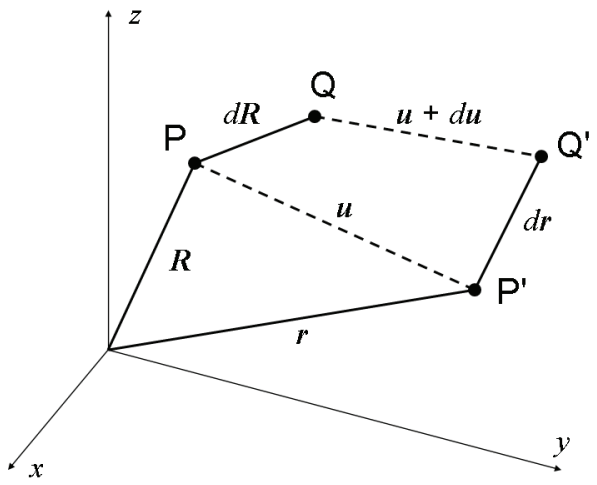


FIGURE 15. Definition of displacement vector $u(R, t)$.

point **P** (position vector $\boldsymbol{R}$) of the body is associated with coordinates X, Y, Z where capital letters are used to denote the reference configuration (a generalization to other rectangular coordinate systems is performed later)

$$\mathbf{P}: \quad (\boldsymbol{R}) = \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = (X, Y, Z)^{\mathsf{T}} \tag{42}$$

Let **Q** with position vector $\boldsymbol{R} + d\boldsymbol{R}$ be a second material point, located at an infinitesimal distance, $(d\boldsymbol{R}) = (dX, dY, dZ)$, from **P**:

$$\mathbf{Q}: \quad (\boldsymbol{R} + d\boldsymbol{R}) = \begin{pmatrix} X + dX \\ Y + dY \\ Z + dZ \end{pmatrix} \tag{43}$$

Under the action of forces, the points **P** and **Q** will be displaced and the body deformed. At some specific point in time, **P** and **Q** are located at new (spatial) positions **P'**, **Q'**, respectively, with associated position vectors (Fig. 15).

$$\mathbf{P'}: \quad (\boldsymbol{R} + \boldsymbol{u}) = \begin{pmatrix} X + u_x \\ Y + u_y \\ Z + u_z \end{pmatrix} = (\boldsymbol{r}) = \begin{pmatrix} x(X,Y,Z) \\ y(X,Y,Z) \\ z(X,Y,Z) \end{pmatrix} \tag{44}$$

$$\mathbf{Q'}: \quad (\boldsymbol{R} + d\boldsymbol{R} + \boldsymbol{u} + d\boldsymbol{u}) = \begin{pmatrix} X + dX + u_x + du_x \\ Y + dY + u_y + du_y \\ Z + dZ + u_z + du_z \end{pmatrix} = (\boldsymbol{r} + d\boldsymbol{r}) = \begin{pmatrix} x + dx \\ y + dy \\ z + dz \end{pmatrix} \tag{45}$$

The vector $\boldsymbol{u}$, in components $(\boldsymbol{u}) = (u_x(X, Y, Z), u_y(X, Y, Z), u_z(X, Y, Z))^{\mathsf{T}}$ is denoted as displacement vector and lowercase letters are used to refer to actual positions. Because $\boldsymbol{r} = \boldsymbol{R} + \boldsymbol{u}$ (see Equation 44), $d\boldsymbol{r}$ can be written as

$$(d\boldsymbol{r}) = \begin{pmatrix} dx \\ dy \\ dz \end{pmatrix}, \text{ whereby}$$

$$\begin{aligned} dx &= \left(\frac{\partial x}{\partial X}\right)dX + \left(\frac{\partial x}{\partial Y}\right)dY + \left(\frac{\partial x}{\partial Z}\right)dZ \\ &= dX + du_x = \left(1 + \left[\frac{\partial u_x}{\partial X}\right]\right)dX + \left(1 + \left[\frac{\partial u_y}{\partial Y}\right]\right)dY + \left(1 + \left[\frac{\partial u_z}{\partial Z}\right]\right)dZ \end{aligned} \tag{46}$$

$$dy = \ldots$$
$$dz = \ldots$$

(dy and dz are obtained by cyclic permutation). Accordingly, in matrix notation

$$(d\boldsymbol{r}) = \left(\left[\boldsymbol{I} + \boldsymbol{J}\right], d\boldsymbol{R}\right) = \left(\boldsymbol{F}, d\boldsymbol{R}\right) \tag{47}$$

The displacement gradient, $\boldsymbol{J}$, is defined as (in components)

$$(\boldsymbol{J}) = \begin{pmatrix} \frac{\partial u_x}{\partial X} & \frac{\partial u_x}{\partial Y} & \frac{\partial u_x}{\partial Z} \\ \frac{\partial u_y}{\partial X} & \frac{\partial u_y}{\partial Y} & \frac{\partial u_y}{\partial Z} \\ \frac{\partial u_z}{\partial X} & \frac{\partial u_z}{\partial Y} & \frac{\partial u_z}{\partial Z} \end{pmatrix} \tag{48}$$

where $\boldsymbol{F} = \boldsymbol{I} + \boldsymbol{J}$ is referred to as the deformation gradient.

In order to obtain a measure for the local strain resulting from a deformation of the body, we calculate the length of $d\boldsymbol{R}$ in the deformed state, $d\boldsymbol{r}$, using

$$|d\boldsymbol{r}|^2 = (d\boldsymbol{r}, d\boldsymbol{r}) = dr^2 = (dx, dy, dz)\begin{pmatrix} dx \\ dy \\ dz \end{pmatrix} = (d\boldsymbol{r})^{\mathsf{T}}(d\boldsymbol{r}) \tag{49}$$

Accordingly, the change in length is

$$\begin{aligned} dr^2 - dR^2 &= \left(\left[\boldsymbol{I} + \boldsymbol{J}\right], d\boldsymbol{R}\right)^{\mathsf{T}}\left(\left[\boldsymbol{I} + \boldsymbol{J}\right], d\boldsymbol{R}\right) - dR^2 \\ &= (d\boldsymbol{R})^{\mathsf{T}}\left[(\boldsymbol{J})^{\mathsf{T}} + (\boldsymbol{J}) + (\boldsymbol{J})(\boldsymbol{J})^{\mathsf{T}}\right](d\boldsymbol{R}) = (d\boldsymbol{R})^{\mathsf{T}}\left[(\boldsymbol{F})^{\mathsf{T}}(\boldsymbol{F}) - (\boldsymbol{I})\right](d\boldsymbol{R}) \end{aligned} \tag{50}$$

where it should be kept in mind that the reference configuration may already have internal strains associated with corresponding stresses.

In the case of small deformations and in the absence of rigid rotations, the nonlinear term, $(\boldsymbol{J})(\boldsymbol{J})^{\mathsf{T}}$, can be neglected and the engineering strain tensor is obtained

$$\frac{1}{2}\left[(J)^{\mathsf{T}} + (J)\right] = \begin{pmatrix} \frac{\partial u_x}{\partial X} & \frac{1}{2}\left(\frac{\partial u_x}{\partial Y}+\frac{\partial u_y}{\partial X}\right) & \frac{1}{2}\left(\frac{\partial u_x}{\partial Z}+\frac{\partial u_z}{\partial X}\right) \\ \frac{1}{2}\left(\frac{\partial u_y}{\partial X}+\frac{\partial u_x}{\partial Y}\right) & \frac{\partial u_y}{\partial Y} & \frac{1}{2}\left(\frac{\partial u_y}{\partial Z}+\frac{\partial u_z}{\partial Y}\right) \\ \frac{1}{2}\left(\frac{\partial u_z}{\partial X}+\frac{\partial u_x}{\partial Z}\right) & \frac{1}{2}\left(\frac{\partial u_z}{\partial Y}+\frac{\partial u_y}{\partial Z}\right) & \frac{\partial u_z}{\partial Z} \end{pmatrix} \quad (51)$$

(The factor ½ derives from the Taylor expansion of the square root, if the actual length, and not the length squared, is calculated).

From Equation 50 it can be seen that the quantity

$$(E) = \frac{1}{2}\left[(F)^{\mathsf{T}}(F) - (I)\right] \quad (52)$$

which is denoted as the Green-Lagrange tensor, also describes strain in the case of large deformations. Because it is as such not influenced by rigid displacements (in contrast to the engineering strain tensor), it is well suited for use in constitutive equations and in nonlinear applications. In addition, the right Cauchy-Green deformation tensor,

$$(C) = (F)^{\mathsf{T}}(F) \quad (53)$$

is likewise often used in constitutive equations. The relation between the displacement vector, $\boldsymbol{u}$, and the deformation gradient, $\boldsymbol{F}$, is sometimes denoted as compatibility condition.

B. Forces: Stress

The previous paragraph is related primarily to solids, because fluids (other than fluids with memory or liquid crystals, which are not taken into account here) have no reference configuration in the sense used before. However, the concept of stress applies equally for fluids and solids. To derive the stress tensor, we consider the interior of a fluid or solid continuum, notably an infinitesimal element *df* at point **P** of an arbitrarily chosen internal "surface" of the body (Fig. 16) obtained by a virtual cut. The continuum may therefore be in a deformed and/or displaced state. A local Cartesian coordinate system is introduced in **P** such that $df = dy\,dz$ and the outer normal (unit vector $\boldsymbol{n}$) of this element has x – direction. We imagine now that one side of the virtually cut continuum be removed. In order that the remaining part of the continuum remains in its state of deformation, the internal forces acting locally before

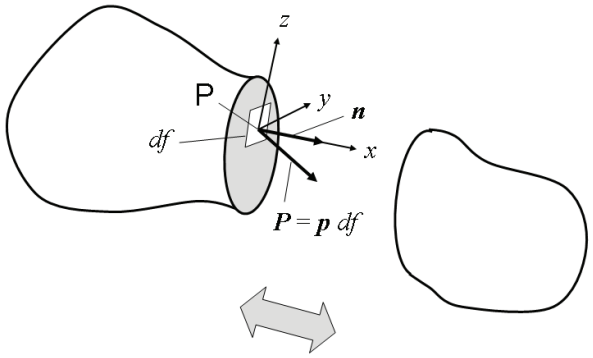


FIGURE 16. Virtual cut through a deformable body.

cutting in the interior of the body along the surface of the virtual cut have to be introduced and applied to the surface. The element of force, $d\boldsymbol{P}$, which the adjacent material, that is, the part of the body that has been removed, exerts locally on *df* is written as $d\boldsymbol{P} = \boldsymbol{p}\,df$ with the aid of the stress vector $\boldsymbol{p}$ (force per unit area). $\boldsymbol{p}$ is now decomposed into the three components

$$\begin{aligned} &\sigma_{xx} = \sigma_x && \text{normal stress} \\ &\tau_{xy},\ \tau_{xz} && \text{shear stresses} \end{aligned} \quad (54)$$

where the indices are chosen such that the first index denotes the outer normal of the surface element under consideration, while the second index indicates the direction of the component. Other virtual cuts that contain the point **P** can be made, in particular such that the associated local surface elements have outer normals in the direction y and z, respectively. Upon application of the same procedure to these other coordinate directions, an ensemble of components is obtained as

$$\boldsymbol{\sigma} = \begin{pmatrix} \sigma_x & \tau_{xy} & \tau_{xz} \\ \tau_{yx} & \sigma_y & \tau_{yz} \\ \tau_{zx} & \tau_{zy} & \sigma_z \end{pmatrix} \quad (55)$$

which is denoted as Cauchy stress tensor (Fig. 17). That this entity has in fact the properties of a tensor is seen when the particular coordinate system (x, y, z) is subjected to a rotation; it will be seen that $\boldsymbol{\sigma}$ then transforms as $\boldsymbol{\sigma}' = \boldsymbol{D}^{-1} \cdot \boldsymbol{\sigma} \cdot \boldsymbol{D}$.

The stress tensor is also symmetric. This follows from angular equilibrium (Fig. 18): the moment acting with respect to the center of mass of the cube around the z - axis is due to the shear stresses τ_{xy} and

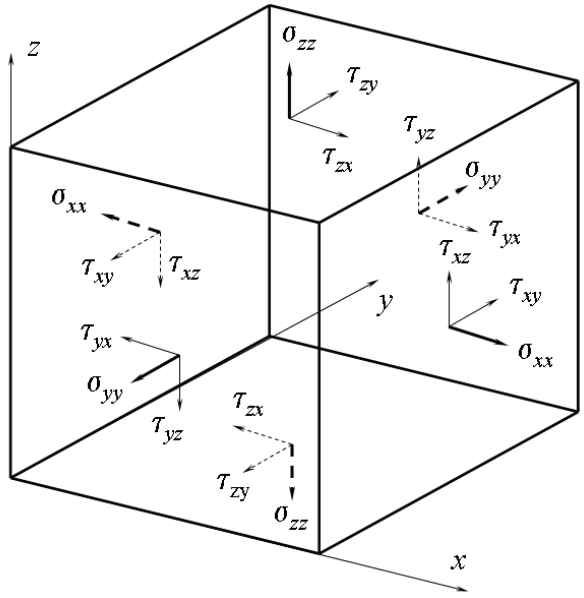


FIGURE 17. Definition of normal and shear stresses.

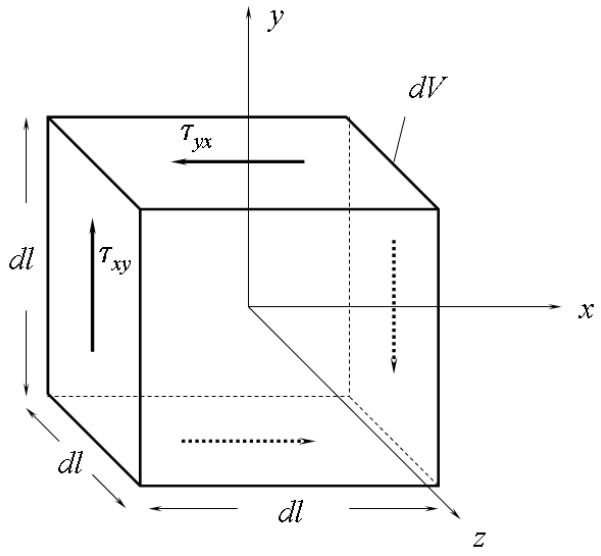


FIGURE 18. Angular momentum balance.

τ_{yx} only and is proportional to the edge length $(dl)^3$, because the forces are proportional to $(dl)^2$ and the moment arm to $(dl)^1$, while the moment of inertia is proportional to $(dl)^5$ because the mass is proportional to $(dl)^3$ and the radius of inertia to $(dl)^2$. If the moment would not be zero (i.e., $\tau_{xy} \neq \tau_{yx}$), there would be infinite rotational accelerations as $dl \rightarrow 0$. The same holds for the other axes, and accordingly the tensor is symmetric. This implies that the rotational momentum balance does not have to be considered within the framework of the present formulation of continuum mechanics because it is always fulfilled. An element ***df*** of the surface area with an arbitrary spatial orientation is now considered (Fig. 19), in components,

$$(\boldsymbol{n})df = (n_x df, n_y df, n_z df) \tag{56}$$

A tetrahedron formed when the axis of the coordinate system is added as shown in the figure. The forces acting on this infinitesimal tetrahedron have to be in equilibrium according to a similar reason as above: the forces are proportional to $(dl)^2$ while the mass is proportional to $(dl)^3$. If the forces were not in equilibrium, the tetrahedron would experience an infinite acceleration as $dl \rightarrow 0$. Accordingly,

$$\begin{aligned} \sigma_x n_x df + \tau_{yx} n_y df + \tau_{zx} n_z df &= p_x df \\ \tau_{xy} n_x df + \sigma_y n_y df + \tau_{zy} n_z df &= p_y df \\ \tau_{xz} n_x df + \tau_{yz} n_y df + \sigma_z n_z df &= p_z df \end{aligned} \tag{57}$$

or,

$$(d\boldsymbol{P}) = (\boldsymbol{p})df = (\boldsymbol{\sigma}, \boldsymbol{n})df \tag{58}$$

Note: Upon application of a transformation ***D*** to Equation 58, one obtains

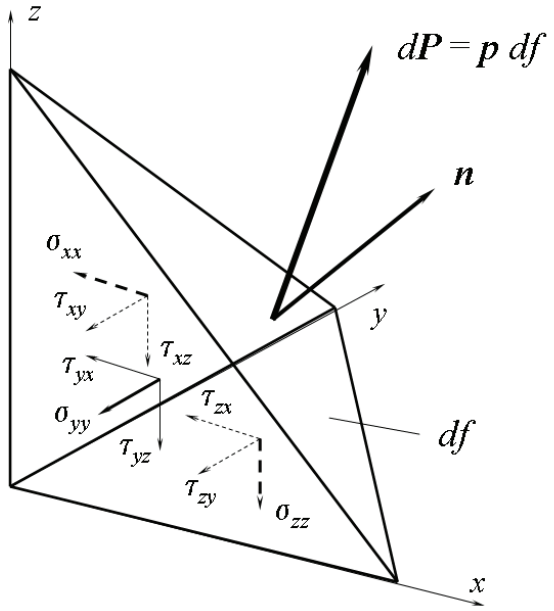


FIGURE 19. Force balance in an infinitesimal tetrahedron.

$$\boldsymbol{D}\cdot d\boldsymbol{P} \;=\; \boldsymbol{D}\cdot\boldsymbol{\sigma}\cdot\boldsymbol{n}\,df \;=\; \boldsymbol{D}\cdot\boldsymbol{\sigma}\cdot\boldsymbol{D}^{-1}\cdot\boldsymbol{D}\cdot\boldsymbol{n}\,df \qquad (59)$$

which shows that $\boldsymbol{\sigma}$ transforms as a tensor.

A clear distinction between normal and shear stresses is important in biomechanics because typical biological materials are fluid-filled. Under the pressures that prevail under physiological conditions, such materials can be considered to be incompressible incompressible (wave propagation phenomena are thereby not taken into account). Accordingly, normal pressures cause in general little biological activity, in contrast to shear stresses, which are mostly responsible for processes involving mechano-transduction, such as the transduction of mechanical into biological signals. When, for example, endothelial cells are exposed to viscous shear stresses caused by blood flow (see below) mechano-transduction occurs due to the deformation of transmembrane molecules. This manifests itself in the up- and down-regulation of gene expression (Fig. 20).

For what follows, solids and fluids have to be treated separately.

1. Solids

The Cauchy stress tensor is always related to a unit surface element regardless of the state of deformation. In the case of large deformations, a surface element given in a specific state of deformation changes its shape and size as the deformation process continues. The Cauchy stress tensor is therefore defined

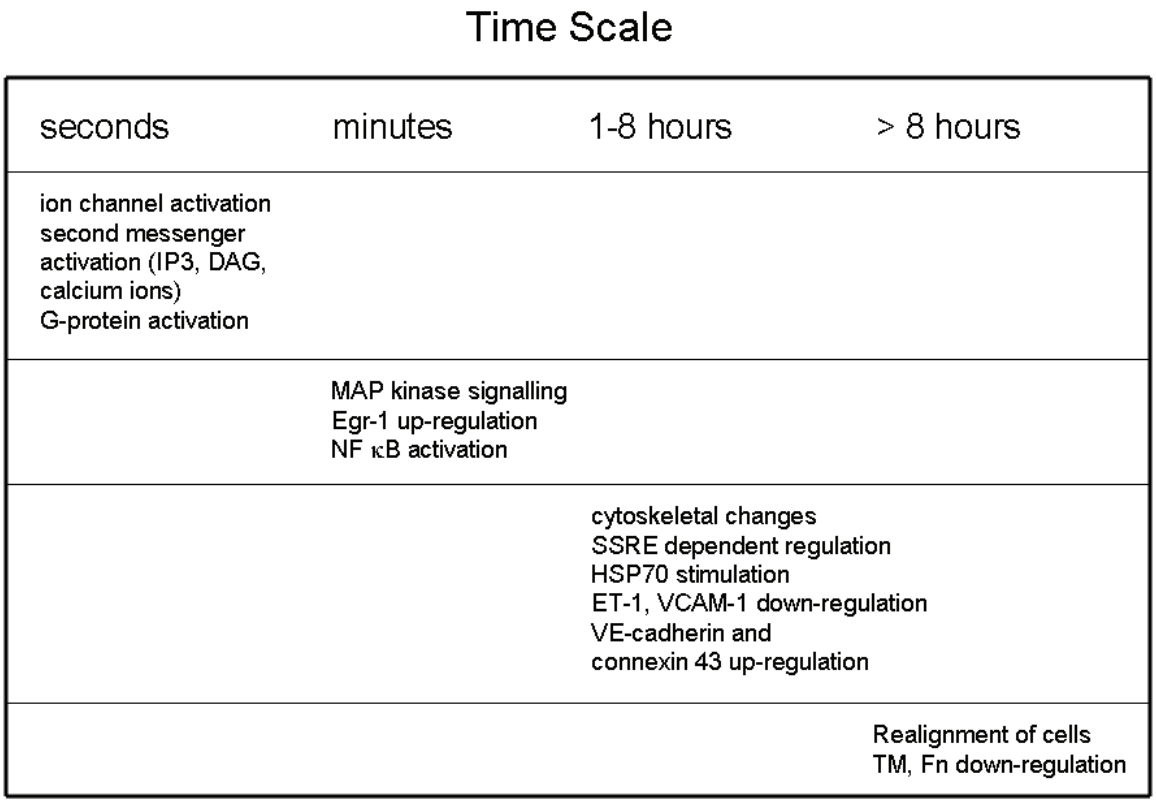


FIGURE 20. Selected cellular reactions to shear stress. Responses are observed within seconds (IP_3, inositol 1,4,5 triphosphate; DAG, diacetylglycerol), minutes (MAP, mitogen-activated protein; Egr-1, early growth response protein 1; NF-κB, nuclear factor κB), hours (SSRE, shear stress response element; HSP70, heat shock protein 70; ET-1, endothelin-1; VCAM-1, vascular cell adhesion molecule; VE-cadherin, vascular endothelial cadherin), and more than 8 hours (TM, thrombomodulin; Fn, fibronectin). (Adapted with permission from Braddock M, Schwachtgen J-L, Houston P, Dickson MC, Lee MJ, Campbell CJ. Fluid shear stress modulation of gene expression in endothelial cells. News Physiol Sci. 1998;13:241–6.)

with respect to a different surface element after each deformation step. A consistent procedure is reached if the stress tensor is always related to a surface element in the reference configuration. This can be obtained in the following fashion. The quantity (***a***, (***b*** × ***c***)) is equal to the volume of a body formed by the three arbitrary, but not collinear vectors ***a***, ***b***, ***c***. For every matrix (***M***) whose determinant is not zero, the following theorem holds (from linear algebra, the formula is related to the volume of a parallelepiped with edges ***a***, ***b***, ***c***).

$$(\boldsymbol{M}\,\boldsymbol{a})^{\mathsf{T}}\,(\boldsymbol{M}\,\boldsymbol{b}\times\boldsymbol{M}\,\boldsymbol{c}) \;=\; \det(\boldsymbol{M})(\boldsymbol{a},\,(\boldsymbol{b}\times\boldsymbol{c}) \tag{60}$$

From this follows

$$(df)(\textit{actual configuration}) \;=\; \det(\boldsymbol{F})(\boldsymbol{F})^{\mathsf{T}}(df)(\textit{reference configuration}) \tag{61}$$

using the deformation gradient, ***F***.

Upon application of this relation to the Cauchy stress tensor, one finds that

$$(\boldsymbol{P}) \;=\; det(\boldsymbol{F})(\sigma)(\boldsymbol{F})^{\mathsf{T}} \tag{62}$$

which is denoted as nominal or first Piola-Kirchhoff stress tensor. It relates the stresses in the actual configuration to the equivalent surface element in the reference configuration. However, this tensor is in general not symmetric, which makes it unpractical for applications. It is therefore convenient to also express the stresses with respect to the reference configuration, which can be achieved by multiplication with $\boldsymbol{F}^{-1}$:

$$(\boldsymbol{T}) \;=\; det(\boldsymbol{F})(\boldsymbol{F}^{-1})(\sigma)(\boldsymbol{F})^{-\mathsf{T}} \tag{63}$$

This is called the second Piola-Kirchhoff stress tensor, and is a useful formulation to describe stresses in large deformation problems. It is again symmetric but it has no direct physical interpretation.

2. Fluids

Fluids other than liquid crystals and without memory effects (the only fluids considered here) have no reference configuration because fluid particles can be rearranged arbitrarily without change in internal force distribution; accordingly, stresses depend only on the momentary global shape as well as on the state of the flow field. They occur, for example, as hydrostatic pressure or shear stresses when internal (viscous, see below) friction effects due to gradients in flow fields are present. Furthermore, biological fluids can generally be considered incompressible under physiological pressures. Incompressibility will therefore be applied throughout the following paragraphs.

C. Constitutive Equations

As mentioned earlier, the mechanical properties of a solid or fluid body with respect to deformability are introduced by way of a constitutive equation that relates the state of strain (including possibly its derivative with respect to time) to the stress state. Solids and fluids again have to be treated differently. Because biological tissues are fluid-filled as a rule, mixture theory is often applied.[94,95]

1. Solids, Hyperelasticity

There are numerous ways to describe the constitutive behavior of the vast variety of solid materials, which can be found in the literature on continuum mechanics. For biological materials, the book of G.A. Holzapfel and R.W. Ogden[96] can be consulted. The seminal work of H. Yamada[97] also has to be mentioned in this context.

A useful method to formulate a constitutive equation that is applicable for large-deformation and nonlinear cases consists of the choice of an appropriate scalar function *W* that describes the elastic strain energy density (hyperelastic material). A scalar function as such is invariant under transformations, and therefore the isotropy group of the material (isotropic, orthotropic, etc.) can be integrated in a straightforward manner. Dissipative effects (internal friction) are thereby not included, however, and have to be added separately.

In order to arrive at a formulation that is valid for nonlinear and large-deformation problems, the second Piola-Kirchhoff stress tensor has to be used, and is obtained according to (in components)

$$T_{ij} \;=\; 2\frac{\partial W}{\partial C_{ij}} \tag{64}$$

This relation is based on the fact that the second Piola-Kirchhoff stress tensor and the Cauchy-Green deformation tensor are connected in the sense that the work per unit of time performed by the stresses equals $\sum_{i,j}(\boldsymbol{T})_{ij}\cdot(\dot{\boldsymbol{C}})_{ji}$ (see, for example, the book of M. Narasimhan[93]). In order that W be invariant, it has to be a function of the invariants I_i associated with $\boldsymbol{C}$. The number of invariants therefore depends on the isotropy group of the material under consideration: three for isotropic, five for transversely anisotropic, etc.

For isotropic materials, the invariants are found to be

$$\begin{aligned} I_1 &= tr(\boldsymbol{C}) \\ I_2 &= \frac{1}{2}\left[\{tr(\boldsymbol{C})\}^2 - tr(\boldsymbol{C}^2)\right] \\ I_3 &= det(\boldsymbol{C}) \end{aligned} \tag{65}$$

For Veronda-Westmann materials, for example, the expression is

$$\begin{aligned} W &= \alpha_1\left(e^{\alpha_2(I_3-3)}-1\right)+\alpha_3(I_2-3) \\ I_3 &= 1 \quad (incompressibility) \end{aligned} \tag{66}$$

α_1, α_2, α_3 are constants that have to be fitted to experimental measurements. Figure 21 shows the application of this expression to uterine tissue of a rabbit in uniaxial tension.

In the case of transverse anisotropy (direction $\boldsymbol{n}$, $|\boldsymbol{n}| = 1$), two further invariants are

$$\begin{aligned} I_4 &= \boldsymbol{n}\cdot\boldsymbol{C}\cdot\boldsymbol{n} \\ I_5 &= \boldsymbol{n}\cdot\boldsymbol{C}^2\cdot\boldsymbol{n} \end{aligned} \tag{67}$$

A model of the myocardium[98] in which such an approximation has been applied is seen in Figure 22.

It should be noted that the more parameters a constitutive relation contains, the more difficult is their experimental determination. The influence of the various parameters may in particular be enormously different, and largely different parameter sets may lead to comparable agreement with measured deformation curves. This is the case especially in highly nonlinear materials and makes it sometimes difficult to compare results of different experiments.

2. Viscoelasticity, Damping

Most biological materials exhibit viscoelastic damping properties that are similar to those associated with "structural" damping, a term that has been adopted from aeroelasticity. This type of damping occurs typically in materials or objects that exhibit a highly inhomogeneous composition. A simple mechanical loading arrangement lends itself to highlight the main feature of structural damping. Assume a strip of material loaded and extended uniaxially

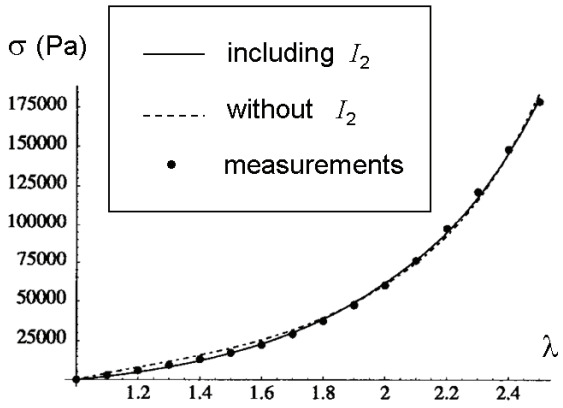


FIGURE 21. Typical stress (σ) – extension (λ) curve for uterine material (rabbit).

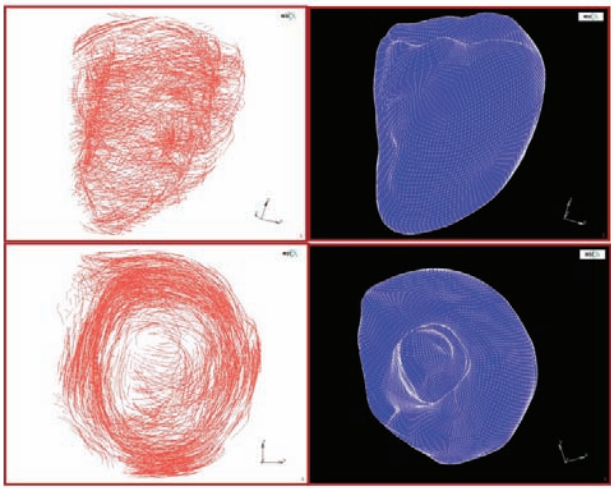


FIGURE 22. Anisotropic finite element model (see section G. Numerical Procedures) of the human heart. The constitutive relation is based on the fiber architecture obtained from peeling a post mortem healthy human heart. A unit vector field was used to describe the anisotropy given by the fiber field. Systole was modeled by adding stepwise contributions to the second Piola-Kirchhoff stress tensor (For further details, the reader is referred to Dorri F, Niederer P, Lunkenheimer PP. A finite element model of the left ventricular systole. Comp Meth Biomech Biomed Eng. 2006;9:319–41.)

by a harmonically oscillating force. Internal friction causes energy to be dissipated. The particular property of structural damping is that the amount of energy dissipated per loading cycle is within a large range independent of the frequency; that is, under a sinusoidal loading cycle, the amount of energy dissipated does—within certain, material-dependent limits—only weakly depend on the speed with which the loading cycle is executed. A consequence of this property manifests itself, for example, in arterial wave propagation: A harmonic wave can be induced in a blood vessel by local external stimulation. As the wave propagates, its amplitude will decay exponentially due to viscoelastic damping. Therefore, the logarithmic decrement of the amplitude of the propagating wave is constant and does not depend on the frequency (verified experimentally between 40 to 200 Hz).

Whenever a linear approximation is possible, structural damping can be implemented by introducing a complex modulus of elasticity, $E = E_1 + iE_2$. To show this, a harmonic oscillator is modeled with this formulation (utilizing a complex spring constant):

$$m\ddot{x} = -(k_1 + ik_2)x \tag{68}$$

The solution reads, upon linearization, with initial amplitude a_0 and frequency parameter $\omega = \sqrt{k_1/m}$,

$$x(t) = a_0 e^{-\frac{k_2}{2k_1}\omega t} \sin(\omega t) \tag{69}$$

which has the desired property because the logarithmic decrement per cycle ($\omega t = 1$) is constant providing that the constitutive parameters (k_1, k_2) are kept constant.

In the case of nonlinear constitutive properties, this complex representation cannot be used. Viscoelastic damping can then be taken into account by including a hereditary integral over the past history of the deformation tensor or its derivative with respect to time. For simplicity, a uniaxial model is considered. A sudden loading of a one-dimensional strip of material with the stress $\sigma_0 \cdot \theta(t)$ ($\theta(t)$ denotes the Heavyside step function, $\theta(t) = \begin{cases} 0 \text{ for } t<0 \\ 1 \text{ for } t \geq 0 \end{cases}$) will lead to a creep process of the elongation

$$\varepsilon(t) = \sigma_0 I(t) \tag{70}$$

with the creep function $I(t)$. Upon generalization to a continuous loading by accumulation of infinitesimal steps $d\varepsilon(t) = d\sigma' I(t - t')$ one obtains

$$\varepsilon(t) = \int_0^t I(t - t') \left(\frac{d\sigma}{dt}\right)_{t=t'} dt' \tag{71}$$

or, after partial integration,

$$\varepsilon(t) = I(0)\sigma(t) + \int_0^t \sigma(t - \tau) \frac{dI(\tau)}{d\tau} d\tau \tag{72}$$

The term $I(0)$ represents the instantaneous elastic response of the material. Notes: (i) A corresponding equation can be derived for the inverse relation (stress as function of the elongation). Instead of the creep function, a function denoted as relaxation function is then defined. (ii) The above derivation is linear. However, a nonlinear elastic behavior can be introduced for $\sigma(t)$. The reader is referred to the literature for more details.[99]

Viscoelasticity manifests itself significantly from the subcellular to the macroscopic level, even in cases in which, due to mostly small deformations such as in bone,[100] linear approximations are useful. It can in particular not be disregarded when the behavior of cells or of the extracellular matrix is analyzed.[101-104] Figure 23 exhibits the instantaneous elastic response and creep of an adherent cell surface.[105]

3. Fluids

The stress state of an ideal, inviscid, isotropic, and incompressible fluid is described by an isotropic tensor that is invariant under transformations (such that no shear stresses occur). Such a tensor has the general form (the minus sign is by definition)

$$(\boldsymbol{\sigma}) = \begin{pmatrix} -p & 0 & 0 \\ 0 & -p & 0 \\ 0 & 0 & -p \end{pmatrix} \tag{73}$$

and describes a hydrostatic pressure state (pressure p).

In the case of an incompressible viscous fluid, stresses develop because of internal friction effects that become effective as soon as the flow field is such that fluid particles execute relative shearing motions. Accordingly, the gradient of the flow field $\boldsymbol{v}(\boldsymbol{r}, \boldsymbol{t})$ has to be considered. Because derivatives

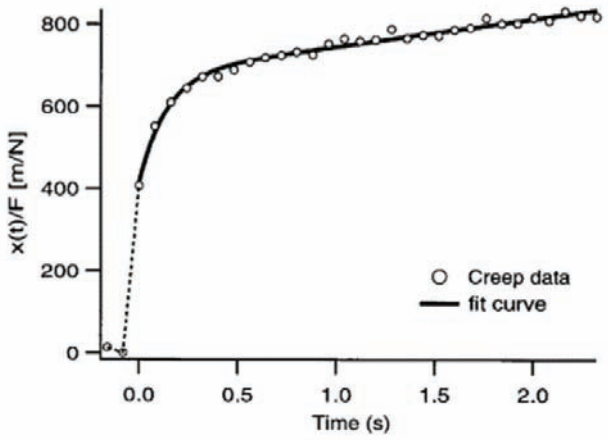


FIGURE 23. Creep behavior of an adherent fibroblast subjected to rapid extension. The force was applied through magnetic beads attached to integrin receptors. The dashed portion of the curve corresponds to the instantaneous elastic response, creep follows. (Used with permission from Bausch AR, Ziemann F, Boulbitch AA, Jacobson K, Sackmann E. Local measurements of viscoelastic parameters of adherent cell surfaces by magnetic bead microrheometry. Biophys J. 1998;75:2038–49.)

only involve infinitesimal changes of a momentary configuration, large displacements and deformations can be disregarded, such that the time derivative of the engineering strain tensor (III.11) can be used to formulate a constitutive law. In the simplest (linear) case, the stress tensor and the velocity gradient (symmetrized) are proportional:

$$(\boldsymbol{\sigma}) = 2\eta\left(\nabla \boldsymbol{v}(\boldsymbol{r},t)_{sym}\right) = 2\eta \begin{pmatrix} \frac{\partial v_x}{\partial x} & \frac{1}{2}\left(\frac{\partial v_x}{\partial y}+\frac{\partial v_y}{\partial x}\right) & \frac{1}{2}\left(\frac{\partial v_x}{\partial z}+\frac{\partial v_z}{\partial x}\right) \\ \frac{1}{2}\left(\frac{\partial v_y}{\partial x}+\frac{\partial v_x}{\partial y}\right) & \frac{\partial v_y}{\partial y} & \frac{1}{2}\left(\frac{\partial v_y}{\partial z}+\frac{\partial v_z}{\partial y}\right) \\ \frac{1}{2}\left(\frac{\partial v_z}{\partial x}+\frac{\partial v_x}{\partial z}\right) & \frac{1}{2}\left(\frac{\partial v_z}{\partial y}+\frac{\partial v_y}{\partial z}\right) & \frac{\partial v_z}{\partial z} \end{pmatrix} \quad (74)$$

The proportionality factor η is denoted as viscosity. Fluids that obey this law as an additive to Equation 73 are called Newtonian fluids. Non-Newtonian fluids such as blood under stationary flow conditions also involve nonlinear expressions in the velocity gradient. Shear stresses due to fluid flow acting on cellular surfaces, in particular in blood vessels, play an important role in mechanotransduction.[106]

D. Equations of Motion

We consider a deformable body (volume *V*) under the influence of forces (Fig. 24). Two types of forces can thereby be distinguished: field forces acting on the entire body without direct contact, such as gravity or electromagnetic forces, and forces due to all kinds of contacts along the surface *S*. Momentum balance then requires

$$\frac{d}{dt}\left[\int_V \rho\, \boldsymbol{v}\, dV\right] = \int_V \rho\, \boldsymbol{k}\, dV + \oint_S \boldsymbol{\sigma}\cdot\boldsymbol{n}\, dS \quad (75)$$

In this equation, ρ denotes the density of the material, $\boldsymbol{v}(\boldsymbol{r}, t)$ the velocity field, $\boldsymbol{k}$ the field force per mass, $\boldsymbol{\sigma}$ the Cauchy stress tensor, and $\boldsymbol{n}$ the outer normal on the surface. Using Gauss' theorem, the last term in Equation 75 can be converted into a volume integral,

$$\oint_S (\boldsymbol{\sigma}, \boldsymbol{n})\, dS = \int_V (\nabla, \boldsymbol{\sigma})\, dV \quad (76)$$

The volume (Fig. 24) can be chosen arbitrarily; that is, the momentum balance has to be fulfilled for any partial volume. Therefore, the equation is valid also for the integrands.

$$\frac{d[\rho\, \boldsymbol{v}]}{dt} = \rho\, \boldsymbol{k} + \nabla\cdot\boldsymbol{\sigma} \quad (77)$$

In this form, the equation appears in what is called the "Lagrangian" formulation, which implies that the

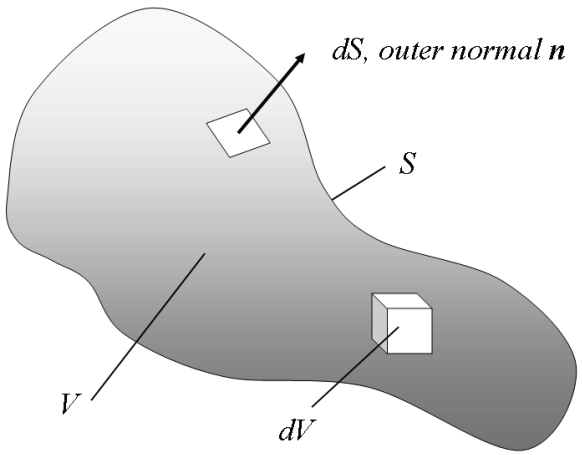


FIGURE 24. Deformable body (volume *V*, surface *S*) considered for the momentum balance.

time derivative is total and follows the material; that is, the observer is fixed to a material point. When soft particle methods (see below) are used for numerical solution purposes, the Lagrangian form of the equation is often applied. It can be converted into the "Eulerian" form by introducing partial derivatives:

$$\frac{\partial(\rho \boldsymbol{v})}{\partial t} + (\boldsymbol{v},\nabla)(\rho \boldsymbol{v}) = \rho(\boldsymbol{k}) + (\nabla,\boldsymbol{\sigma}) \quad (78)$$

In this formulation, the observer is located at a fixed point in space. Due to the second term on the left side of the equation, it contains an unpleasant nonlinear term in the (usually to be determined) quantity $\boldsymbol{v}(\boldsymbol{r}, t)$. Derivatives, however, appear only linearly, which implies that the general theory of linear partial differential equations is still valid (quasilinear equations). Solids and fluids have again to be treated separately.

1. Solids

The velocity field is usually expressed with the aid of the displacement field, such that a second-order partial differential equation is obtained. This equation is then worked out with the constitutive equation of choice that relates deformations with stress such that the equation is determined (along with initial and boundary conditions, depending on the problem at hand). Because there is a wide variety of solid materials with quite different constitutive equations, many different formulations of Equations 77 and 78, respectively, can be found throughout the literature.

2. Fluids

In the case of an incompressible Newtonian fluid, the equation of motion can readily be derived using Equations 73 and 74 and a unique form exists. It is therefore treated in more detail in the following. Inserting Equation 73 and Equation 74 into Equation 78 yields the Navier-Stokes equation

$$\rho \frac{\partial \boldsymbol{v}}{\partial t} + \rho(\boldsymbol{v},\nabla)\boldsymbol{v} = \rho \boldsymbol{k} - \nabla p + \eta \nabla^2 \boldsymbol{v} \quad (79)$$

Incompressibility requires furthermore $(\nabla,\boldsymbol{v}) = 0$. This equation is obtained by balancing the inflow and outflow of a volume element ΔV (see below). The nonlinear term mentioned above usually prevents analytic solutions; numerical procedures have mostly to be applied (computational fluid dynamics). The most important analytic solutions are treated further down.

Note: If the last term in Equation 79 is dropped and only Equation 73 is used (inviscid incompressible fluid), a set of equations called Euler-Bernoulli equations is obtained. Incompressibility also implies that thermodynamic considerations do not have to be taken into account.

Every partial volume of a continuum has an infinite number of degrees of freedom. In the case of homogeneous media, equations of continuum mechanics (fluid or solid) are therefore scalable as a rule. This is achieved by non-dimensionalization of the equations and is demonstrated in the following with Equation 79 for stationary flow and without the $\rho\boldsymbol{k}$ term:

$$\rho(\boldsymbol{v},\nabla)\boldsymbol{v} = -\nabla p + \eta \nabla^2 \boldsymbol{v} \quad (80)$$

We define the dimensionless quantities $\boldsymbol{U} = \frac{\boldsymbol{v}}{U_0}$, $\boldsymbol{R} = \frac{\boldsymbol{r}}{D_0}$, $P = \frac{\rho D_0}{U_0 \eta}$, where U_0 and D_0 are quantities that can to some extent be chosen arbitrarily. They have to be characteristic for the problem under investigation; for example, the flow velocity far away from the object under consideration and the diameter of a tube. Upon substitution of these quantities, one arrives at

$$\frac{\rho U_0 D_0}{\eta}(\boldsymbol{U},\nabla)\boldsymbol{U} = -\nabla P + \nabla^2 \boldsymbol{U} \quad (81)$$

The dimensionless parameter $(\rho U_0 V_0)/\eta$ is called Reynolds number, and serves as a scaling factor. It characterizes the relation between inertial and viscous forces.

The significance of the Reynolds number as a scaling quantity becomes apparent when the transition from laminar to turbulent flow is considered. Turbulence is characterized by a stochastic variation of the flow velocities and is observed when the Reynolds number exceeds some critical value. This implies that the inertial forces dominate over the viscous damping forces such that minute momentary fluctuations or aberrations from the laminar (not necessarily steady)

flow grow exponentially. Problems having the same Reynolds number exhibit similar flow characteristics.

Other scaling factors can be defined for other continuum mechanics problems. A further such scaling quantity will be introduced in the next chapter after a transformation of the Navier-Stokes equation to other coordinate systems.

Many convective transport conduits in biology are long and cylindrical, and axial flow prevails (blood vessels, lymph vessels, urethra, etc.). For such problems, cylindrical coordinates are more useful to describe such systems than Cartesian coordinates. A general transformation of a set of Cartesian coordinates x_1, x_2, x_3 into a curvilinear, orthogonal system ξ_1, ξ_2, ξ_3 by way of the functions

$$\begin{aligned} x_1 &= x_1(\xi_1,\xi_2,\xi_3) \\ x_2 &= x_2(\xi_1,\xi_2,\xi_3) \\ x_3 &= x_3(\xi_1,\xi_2,\xi_3) \end{aligned} \tag{82}$$

has been performed in the Introduction.

Utilizing the formulas derived in the Introduction, one finds, after some calculation, the Navier-Stokes equations in cylindrical coordinates r, ϑ, z, written in components

$$\begin{aligned} \rho\left[\frac{\partial v_r}{\partial t} + (\boldsymbol{v},\nabla v_r) - \frac{v_\vartheta^2}{r}\right] &= -\frac{\partial p}{\partial r} + \eta\left[\nabla^2 v_r - \frac{v_r}{r^2} - \frac{2}{r^2}\frac{\partial v_\vartheta}{\partial \vartheta}\right] \\ \rho\left[\frac{\partial v_\vartheta}{\partial t} + (\boldsymbol{v},\nabla v_\vartheta) - \frac{v_\vartheta v_r}{r}\right] &= -\frac{1}{r}\frac{\partial p}{\partial \vartheta} + \eta\left[\nabla^2 v_\vartheta - \frac{v_\vartheta}{r^2} - \frac{2}{r^2}\frac{\partial v_r}{\partial \vartheta}\right] \\ \rho\left[\frac{\partial v_z}{\partial t} + (\boldsymbol{v},\nabla v_z)\right] &= -\frac{\partial p}{\partial z} + \eta\nabla^2 v_z \end{aligned} \tag{83}$$

along with the equation of continuity, implying incompressibility,

$$\frac{1}{r}\frac{\partial}{\partial r}(r v_r) + \frac{1}{r}\frac{\partial v_\vartheta}{\partial \vartheta} + \frac{\partial v_z}{\partial z} = 0 \tag{84}$$

The latter equation is obtained from a balance of inflow and outflow of an incompressible fluid with respect to a volume element (Fig. 25) with edge lengths $h_i\, d\xi_i$:

Inflow: $v_{\xi_1}(\xi_1,\xi_2,\xi_3)\cdot h_2\, d\xi_2\, h_3\, d\xi_3$ + *cyclic permutation*

Outflow: $v_{\xi_1}(\xi_1+d\xi_1,\xi_2,\xi_3)\cdot h_2\, d\xi_2\, h_3\, d\xi_3$ + *cyclic permutation*

The general concepts outlined above on the dynamics of continua are of paramount significance in biomechanics because motion and deformation are primary features associated with life. Therefore, in addition to the selected references given below,[107-110] most of the literature cited earlier and farther down in this chapter are based on these equations in various formulations. Continuum aspects are also widely applied in biomechanics research relating to animals, in particular fishes,[111-115] insect and bird flight,[116-119] and growing of plants.[120-126]

E. Analytic Solutions of the Navier-Stokes Equations of Use in Biomechanics

In contrast to solids, in which a large variety exists and therefore many different formulations of the continuum mechanical equations can be derived, the Navier-Stokes equations are quite universal (non-Newtonian fluids can also be modeled) and are often used in biomechanics. However, due to their quasi-linear nature (the second-order term is involved in a first-order derivative only), the Navier-Stokes equations have only a few analytical solutions. Some of them are of importance in biomechanics and are presented in the following. In particular, in cases of one-dimensional parallel flow in direction z ($v_r = v_\vartheta = 0$), it follows from Equation 79 and is explicitly seen in Equation 83 that $\partial v_z/\partial z = 0$, such that the nonlinear terms in Equation 78 vanish. Analytic solutions are then possible (sections 1–3, below). A somewhat different situation exists in section 4, where an analytic procedure is feasible in spite of nonlinearity. Inviscid

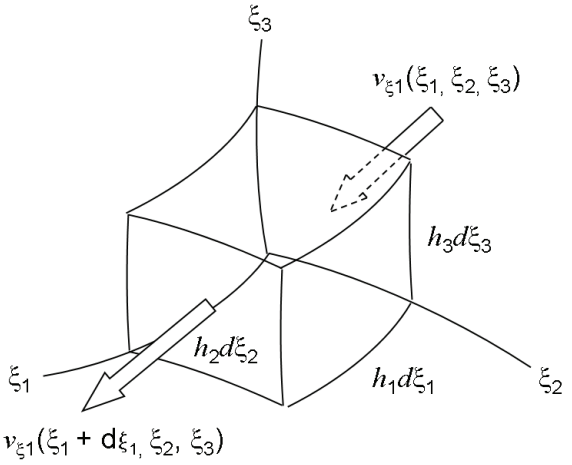


FIGURE 25. Fluid flow balance in an infinitesimal volume element.

flow, as discussed in section 5, allows for yet a different approach, while in some cases, such as in section 6, approximations lead to a useful result.

A common feature of real fluids is that there is no slip at rigid surfaces. This fact is often used in the form of boundary conditions.

1. Hagen-Poiseuille Flow

The well-known Hagen-Poiseuille flow is obtained when one-dimensional, stationary, laminar flow of an incompressible Newtonian fluid in an infinite straight and rigid circular tube (radius R) is considered. With $v_r = v_\vartheta = 0$ and discarding the time derivatives, the three equations in Equation 83 reduce to

$$\begin{aligned} \frac{\partial p}{\partial r} &= \frac{\partial p}{\partial \vartheta} = 0 \\ \frac{d^2 v_z}{dr^2} + \frac{1}{r}\frac{dv_z}{dr} &= \frac{1}{\eta} p_z \end{aligned} \tag{85}$$

From the continuity equation (Equation 84) follows $\frac{\partial v_z}{\partial z} = 0$ such that the nonlinear term vanishes, as mentioned above. Since the problem is translationally invariant, the pressure gradient, $\frac{dp}{dz} = p_z$ is constant, and is assumed to be prescribed. It represents an inhomogeneity for the therefore ordinary differential equation (Equation 85). With the usual procedure for inhomogeneous equations—add a particular solution of the inhomogeneous equation to the general solution of the homogeneous equation, then apply the boundary conditions $v_z\,(r = R) = 0$, $v_z\,(r = 0) < \infty$)—the well-known parabolic velocity profile associated with the Hagen-Poiseuille flow is found

$$v_z = \frac{(-p_z)R^2}{4\eta}\left(1 - \frac{r^2}{R^2}\right) \tag{86}$$

where the total flow (integrated over the cross section) $Q = (-p_z)\,R^4/(8\eta)$.

The critical Reynolds number for tube flow is $Re \approx 2200$, above which inertial forces dominate over friction such that, as mentioned previously, minute aberrations from the stationary flow conditions immediately grow exponentially and turbulence sets in. The transition from laminar to turbulent flow is nontrivial, in particular in non-stationary flow, and is subject to extensive research, also in biofluid mechanics.

2. Witzig-Womersley Flow

We assume the same geometrical and axial flow conditions as above, but the prescribed pressure gradient is now assumed to vary harmonically in time, $\frac{dp}{dz} = p_z^0\, e^{i\omega t}$. Because the equation is linear, the solution for an arbitrary time dependence of the pressure gradient can be obtained by Fourier superposition. With the separation ansatz $v_z(r,t) = v(r)e^{i\omega t}$, the equation is

$$\frac{d^2 v}{dr^2} + \frac{1}{r}\frac{dv}{dr} - \frac{i\rho\omega}{\eta} v = \frac{1}{\eta} p_z^0 \tag{87}$$

where the time dependence cancels is obtained. Upon application of the transformation $\zeta^2 = i^3 \frac{\omega\rho R^2}{\eta}$ the homogeneous equation becomes a zero-order Bessel equation

$$\frac{d^2 v}{d\zeta^2} + \frac{1}{\zeta}\frac{dv}{d\zeta} + v = 0 \tag{88}$$

The particular solution (use $v = const.$) of the inhomogeneous equation and the same boundary conditions as above one arrives at

$$v_z(r,t) = \frac{-p_z^0}{i\rho\omega}\left(1 - \frac{J_0\left(i^{3/2}\,\alpha\,\frac{r}{R}\right)}{J_0\left(i^{3/2}\,\alpha\right)}\right) e^{i\omega t} \tag{89}$$

with the dimensionless Witzig-Womersley parameter $\alpha^2 = \frac{\omega\rho R^2}{\eta}$ (Fig. 26)†.

This parameter, like the Reynolds number, characterizes the ratio of the inertial forces to the viscous forces. While in stationary flow, however, the inertial forces are due to the convective contribution, in oscillating flow the inertial forces are due to the explicit time dependence and are dominant. (This is often also the case when the convective term in the Navier-Stokes equations does not disappear). As is seen from Equation 89, the velocity profile depends uniquely on α.

Because the tube is assumed to be rigid, there is no pulse propagation (formally, the wave speed is infinite), in that all fluid particles along a stream line oscillate in phase. If a deformable tube is modeled,

† This solution was first presented by Konrad Witzig (1914) in his dissertation at the University of Berne, Switzerland.

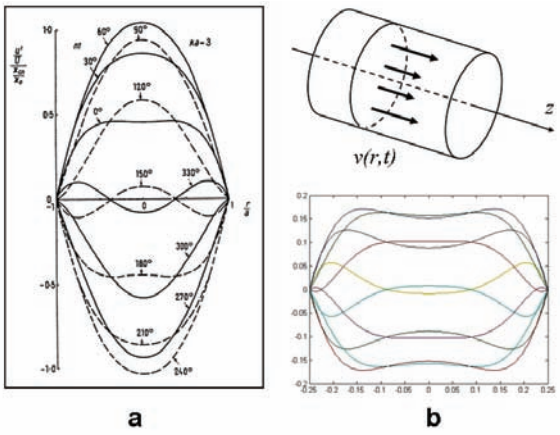


FIGURE 26. Witzig-Womersley flow. a, Original drawing from Witzig's dissertation (Bern 1914) showing typical velocity profiles in 30 deg intervals of a sinusoidal cycle. b, Matlab simulation for α = 6.7.

formulas for the wave speed can be obtained; however, only approximate solutions can be found.

Witzig-Womersley flow is extensively applied in the mammalian cardiopulmonary system, including blood and air flow. Mostly laminar conditions are thereby found; transition to turbulent flow is difficult to observe and a mathematical analysis is formidable. Turbulent flow may be observed in the ascending aorta in cases of aortic valve insufficiency and concurrent regurgitation, in which the heart produces jet flow, and in the lung under conditions of heavy breathing. Note that eddy shedding as occurs, for example, in bifurcations is not necessarily a sign of turbulence although some irregularity exists.

Note: Equation 89 can readily be integrated over the cross-section using theorems on Bessel functions to obtain the oscillating total flow. Furthermore, for $\omega \rightarrow 0$, Hagen Poiseuille flow should be obtained. That this is the case is seen from the power expansion of the zero-order Bessel function

$$J_0 = 1 + i\left(\frac{\alpha^2 r^2}{4R^2}\right) + \ldots \qquad (90)$$

when it is inserted into Equation 89.

3. Laminar Boundary Layer

The concept of "boundary layer" was introduced by Ludwig Prandtl in 1905 and reflects the fact that viscous effects are of particular influence in the neighborhood of surfaces where velocity gradients are high. The simplest case of a laminar boundary layer (there are also turbulent boundary layers) is obtained when a Newtonian fluid over an oscillating plane is considered (Fig. 27). Cartesian coordinates are used here, and the plane is assumed to oscillate according to $(v)=(v_0 \cos \omega t, 0, 0)$. The solution is based on the ansatz $v_y = v_z = 0$, $v_x = v_x(y)$ only, and, furthermore, $\frac{\partial p}{\partial x} = 0$ because the problem is translationally invariant and the pressure is assumed to remain finite at infinity. The Navier-Stokes equations, along with the equation of continuity, yield

$$\frac{\partial p}{\partial y} = \frac{\partial p}{\partial z} = 0$$
$$\rho \frac{\partial v_x}{\partial t} = \eta \frac{\partial^2 v_x}{\partial y^2} \qquad (91)$$

From the solution

$$v_x(y,t) = v_0 e^{-ky} \cos(\omega t - ky), \quad \text{whereby } k^2 = \frac{\omega\rho}{2\eta}, \qquad (92)$$

it is seen that at a distance of $\delta = \sqrt{\frac{2\eta}{\omega\rho}}$ the amplitude of the oscillation decays by a factor of $1/e$. δ is denoted as the thickness of the boundary layer.

4. Couette Flow

This type of flow occurs generally in symmetrically

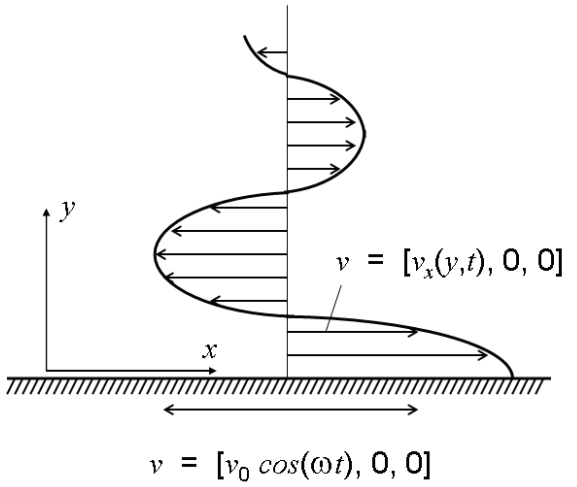


FIGURE 27. The simplest model of a laminar boundary layer.

rotating flow problems. A particular case is of importance, for example, in viscosimeters of the form shown in Figure 28. Note that such viscosimeters provide correct results for the (apparent) viscosity also in the case of non-Newtonian fluids (in contrast to capillary-based viscosimeters that make use of the Hagen-Poiseuille formula valid for Newtonian fluids only). With the ansatz $(v) = (0, v_\vartheta(r), 0)$, $p = p(r)$ and after making use of the continuity equation, one arrives at two equations (note that the convective term does not disappear),

$$\rho\frac{v^2}{r} = \frac{dp}{dr} \quad \text{and} \quad \frac{d^2v}{dr^2} + \frac{1}{r}\frac{dv}{dr} - \frac{v}{r^2} = 0 \tag{93}$$

With the boundary conditions $v(r = R_i) = 0$, $v(r = R_a) = \omega R_a$, the solution reads

$$v(r) = \frac{\omega R_a^2}{R_a^2 - R_i^2}\left(r - R_i^2\frac{1}{r}\right) \tag{94}$$

from which the moment acting on the inner (rotating) cylinder

$$M = 4\pi\eta\omega\, R_a^2 R_i^2 \frac{1}{R_a^2 - R_i^2} \tag{95}$$

is obtained. This allows the determination of the (effective) viscosity; that is, the viscosity valid for the shear field given by ω also for non-Newtonian fluids, because the solution (Equation 94) is independent of the viscosity.

5. Inviscid Flow

When the viscous term in the Navier Stokes equation is dropped, the Euler-Bernoulli equation is obtained, as mentioned previously (an external force term, $\rho\, \boldsymbol{k}$, is also not considered):

$$\rho\frac{\partial \boldsymbol{v}}{\partial t} + \rho(\boldsymbol{v}, \nabla)\boldsymbol{v} = -\nabla p \tag{96}$$

In case of incompressibility and stationary flow (non-turbulent), the well-known relation

$$p + \rho\frac{v^2}{2} = const. \tag{97}$$

is obtained, when the equation is integrated along a stream line. Under flow conditions where eddy shedding and flow separation is of secondary importance only, this formula provides a reasonable approximation, for example, for a flow through a constriction (stenosis) and associated pressure variations.

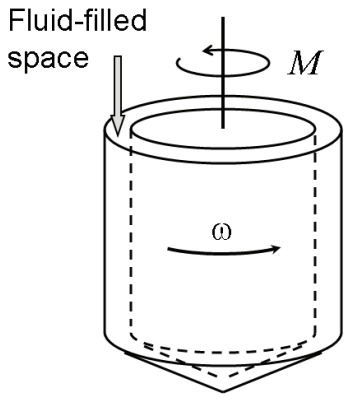


FIGURE 28. Couette flow in a double-cylinder geometry.

F. Hydraulic Approximation

Tube flow occurs ubiquitously in biology; for example: flow in blood vessels, arterial pulse propagation, lymph flow, flow in the urethra, tear channel, and to some extent also air flow in the conduits of the lung. Hagen-Poiseuille and Witzig-Womersley flow can typically be used to model such flow situations. However, only an average axial velocity in the cross-section instead of the true velocity profile is often of interest, particularly because real biological conduits always exhibit some tortuosity and varying cross-sections. A useful approximation can be found for such conditions, called hydraulic approximation.

We start from the Navier-Stokes and continuity equations (Equations 83 and 84). The first assumptions are $v_\vartheta = 0$, $v_r << v_z$ because the main flow is in the axial direction. The first two equations (in Equation 83) imply that the pressure may be assumed to be constant over a given cross-section. In the axial direction (third equation), the equation is treated as follows

$$\frac{\partial v_z}{\partial z} + v_r \frac{\partial v_z}{\partial r} + v_z \frac{\partial v_z}{\partial z} = -\frac{1}{\rho}\frac{\partial p}{\partial z} + \frac{\eta}{\rho}\left(\frac{\partial^2 v_z}{\partial r^2} + \frac{1}{r}\frac{\partial v_z}{\partial r} + \frac{\partial^2 v_z}{\partial r^2}\right)$$

$$\overbrace{\frac{\partial}{\partial r}(v_r v_z) - v_z \frac{\partial v_r}{\partial r}}$$

$$\overbrace{v_z \frac{\partial v_z}{\partial z} + \frac{v_r v_z}{r}} \qquad (98)$$

where the continuity equation has been applied.

The resulting equation is averaged over the cross-section

$$\int_0^{2\pi}\int_0^R \left[\frac{\partial v_z}{\partial z} + \frac{\partial (v_z)^2}{\partial z} + \frac{1}{r}\frac{\partial}{\partial r}(r v_r v_z)\right] r\, dr\, d\vartheta = \int_0^{2\pi}\int_0^R \left[-\frac{1}{\rho}\frac{\partial p}{\partial z} + \frac{\eta}{\rho}\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial v_z}{\partial r}\right)\right] r\, dr\, d\vartheta \qquad (99)$$

After performing the integration, thereby approximating $\int_0^{2\pi}\int_0^R \left[\frac{\partial (v_z)^2}{\partial z}\right] r\, dr\, d\vartheta$ with $\frac{\partial}{\partial z}(A v^2)$, and utilizing the continuity equation, one obtains for the axial velocity v, averaged over the cross-section

$$\frac{\partial v}{\partial t} + v\frac{\partial v}{\partial z} + \frac{1}{\rho}\frac{\partial p}{\partial z} = f \qquad (100)$$

The friction term,

$$f = \int_0^{2\pi}\int_0^R \frac{\eta}{\rho}\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial v_z}{\partial r}\right) r\, dr\, d\vartheta = \cdots \frac{2\pi R}{\rho}\left(\eta\frac{\partial v_z}{\partial r}\right) = \frac{2\pi R}{\rho}\tau_w. \qquad (101)$$

τ_w thereby denotes the skin friction at the wall, according to Newton's friction law (Equation 74). There are various ways (nontrivial, but not worked out here) that this term (which depends on the velocity profile) can be approximated. Furthermore, the term

$$\int_0^{2\pi}\int_0^R \left[\frac{1}{r}\frac{\partial}{\partial r}(r v_r v_z)\right] r\, dr\, d\vartheta \qquad (102)$$

can be integrated and vanishes providing that there is no leakage through the wall.

A similar (straightforward) averaging scheme is used for the continuity equation which yields

$$\frac{\partial}{\partial z}(A v) + \frac{\partial A}{\partial t} = 0 \qquad (103)$$

Together with a constitutive equation for the wall $A[p(z,t),z]$, which contains the anatomical information (explicit dependence on z) and may include viscoelastic contributions, as well as an approximation for the friction term, Equations 100 and 103 represent the hydraulic approximation. This approximation, along with related methods involving one-dimensional wave propagation modeling, lend themselves for use in the analysis of arterial pulse waves[128-135] (Fig. 29).

G. Numerical Procedures

Most realistic problems in biomechanics cannot be solved analytically; accordingly, numerical solution procedures have to be chosen. Nevertheless, it is often advisable that a simplified analytical model is considered before a numerical analysis is performed to obtain a general idea about the properties of the problem under consideration on the one hand, and on the other, to provide a basis for model validation.

While ordinary differential equations can be solved numerically by various standard methods depending on the degree of stiffness, there are two general procedures to solve partial differential equations with the aid of a computer. Either the equations themselves are approximated (i.e., discretized) or the solution. Euler or Lagrange representation can likewise be chosen.

When the partial differential equations are discretized, differential quotients are approximated by finite difference quotients using the finite difference method; that is, the equations themselves are approximated.[136,137] The resulting (in general, large) set of linear algebraic equations is solved within machine precision. However, care has to be exercised in the discretization of the partial derivatives (second-order derivatives, stacking of derivatives of pressure and velocity); appropriate schemes can be found in the literature.

The method of choice for many modeling problems in biomechanics is the finite element method.[138,139] The method was first applied mainly in orthopedics,[140] but since then has enjoyed widespread applications.[141-151] In this method, the equations remain in their original form, but the solution is piecewise approximated. The method is particularly well suited for irregular geometries and nonlinear problems. Figure 30 exhibits the compression of a trabecular bone sample, experiment, and associated

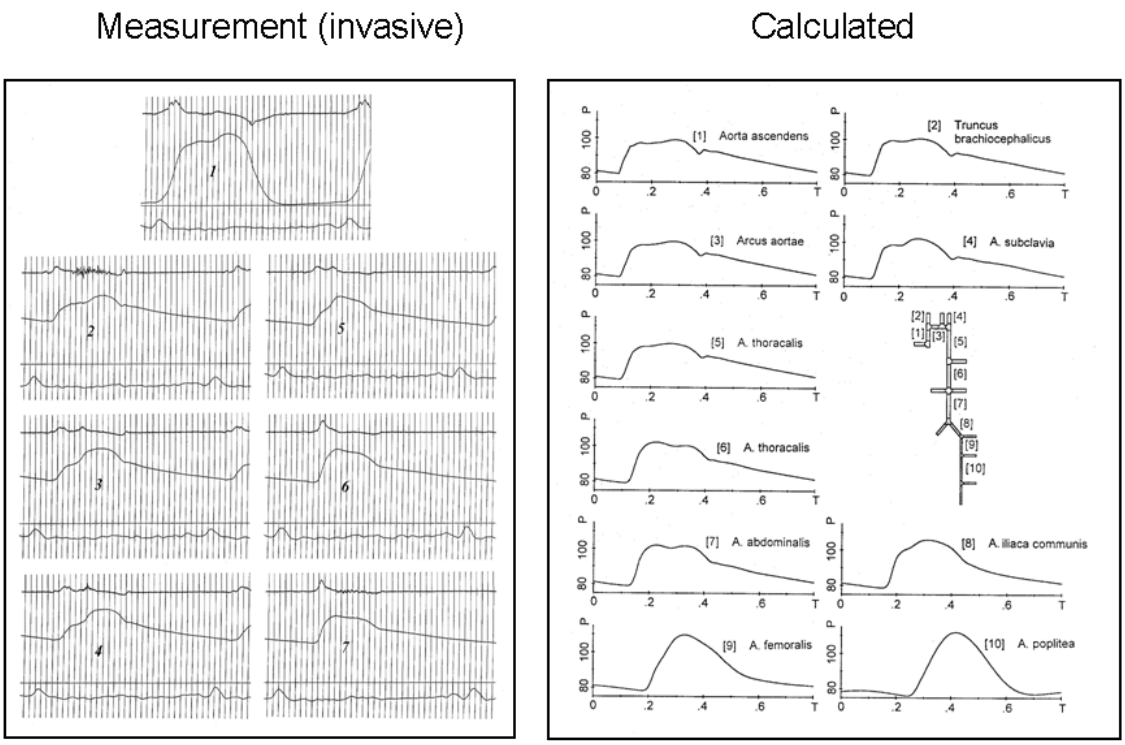


FIGURE 29. Pulse propagation along the aorta. Left: Measurements (invasive); *1*: left ventricle, *2*: ascending aorta, …, *7*: abdominal aorta. Right: Theoretical model based on the hydraulic approximation. (For further details, the reader is referred to Niederer P, Knothe Tate ML, Steck R, Boesiger P. Some remarks on intravascular and extravascular transport and flow dynamics. Int J Cardiovsc Med Sci. 2000;3:21–31.)

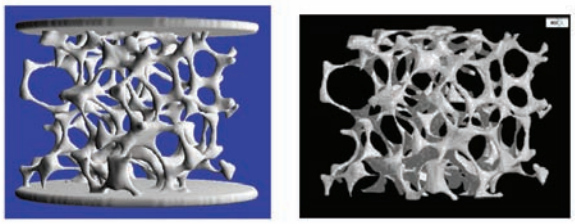


FIGURE 30. Image guided failure analysis. Deformation of a trabecular bone sample under compressive loading as recorded by micro Computed Tomography (μCT) (left), corresponding finite element model (right). (Courtesy R. Müller, ETH Zurich.)

theoretical finite element model. In addition, a number of highly optimized commercial programs are available. Because biological materials are mostly fluid filled, they can be considered as incompressible under the typical pressures occurring under physiological conditions. However, incompressibility may pose substantial problems, particularly in nonlinear, large-deformation finite element models due to "element locking" (element deformation observing incompressibility cannot simply be achieved). Special procedures exist to deal with this problem.

A further drawback of the finite element method is associated with meshing, especially if the problem to be solved requires (time-consuming) re-meshing, as is the case, for example, for surgery simulation. A similar problem exists with finite difference methods. "Mesh-free" methods are therefore of interest. Such a method derives from the soft or smoothed particle method, originally derived in astrophysics. In this method, re-meshing problems can to some degree be avoided. The method consist of incorporating field quantities into radially symmetrical "smoothed particles" whose location and size are largely arbitrary, but have to allow for a suitable representation of the problem including all of the desired details. Particles are created with the aid of a scalar function $\boldsymbol{W}(\boldsymbol{r}, h)$, which is called the smoothing kernel with core radius h, zero outside (finite support).

The function is radially symmetrical and normalized, $\int_{|r|\le h} W(r,h)dr = 1$ such that conservation is maintained. Typical functions have an inverted bell-like form (and are zero for $|r| \ge h$). Any field quantity A (scalar, vectorial, or tensorial such as density, velocity, momentum, etc.) is interpolated in particle form as

$$A_S(r) = \sum_j m_j \frac{A_j}{\rho_j} W(r - r_j) \quad (104)$$

where j iterates over all particles
m_j is the mass of a particle
ρ the density
r_j the position
A_j the field quantity at r_j

The subscript "S" denotes a smoothed quantity. As is seen from Equation 104, derivatives act only on W. For dynamic simulations, the interaction between particles in the form of force fields has to be modeled such that the physical behavior of the material (fluid, solid) is correctly approximated. Often, the equations are formulated and solved in Lagrange form.[152-157] These methods are particularly well suited for applications in which execution time is critical, such as the case of surgery simulation.

IV. DIFFUSION AND OSMOSIS

A. Diffusion

Typical biological tissues are enormously inhomogeneous with respect to their composition; they are also of a multiphase nature and include a variety of fluid and solid phases. Water is therefore the primary component, because about 55% to 70% of the volume of the human body consists of water (a general decrease with age is normal). Accordingly, the transport of nutrients, oxygen, hormones, cytokines, etc., from the location of uptake (intestine, lung) or production (glands, liver) to the cells of the various organs occurs mostly in a water solution; however, transport is modulated further by membranes and gel-like structures. Likewise, waste products are brought to the excreting organs (kidney, liver, lung) along similar paths. Intracellular transport is also to a large extent fluid-based, although transport by or along carrier or directing molecules is dominant.

Driving mechanisms for transportation in biological systems are based on:

- convection
- diffusion
- transport of electrically charged particles due to potential differences
- active transport by or along transport molecules
- pinocytosis
- transport by cells

While diffusion is due to concentration gradients and as such is of a passive nature, convection is driven by pressures that are caused by the action of muscles, gravity, or deformations. Both transport mechanisms can be described systematically by equations such as Fick's law for diffusion (discussed below) and the Navier-Stokes equations for convection (fluids, discussed in the last chapter). Because biological materials contain many constituents that are electrically charged (ions), electric potential difference, for example, across membranes, also modulate transport dynamics. Such processes can also be modeled mathematically and are extensively used in electrophysiology. In contrast, transport by carrier molecules and pinocytosis (vesicle transport through cells) involves active processes that are characterized by a great variety of different mechanisms, most of which defy a "simple" mathematical description. In the case of a single transport molecule, the mass action principle leads to Michaelis-Menton kinetics, where the change of concentration of a substance carried by the molecule as function of time follows the relation $\dot{c} = -(V_m \cdot c)/(k_m + c)$ with constants V_m (maximal transportation capacity when concentrations are large) and k_m (for small concentrations, the transport equation becomes linear). Transportation by cells occurs; for example, in the case of red blood cells that carry oxygen and CO_2 or of antigen-presenting leucocytes. Again, such transportation modalities cannot be modeled mathematically in a straightforward manner.

In general engineering practice, convection is considered rapid, effective, and useful for the transportation of large quantities, while diffusion is

thought to represent a slow process. In biology, however, this notion has to be applied with care because transport distances to be covered by convection and diffusion differ greatly.

Typical convection distances and times, for example, in the systemic circulation, are on the order of centimeters to meters and seconds to minutes, respectively. The intravascular volume (typically 5 L of blood for an adult) circulates at rest about once per minute on average because cardiac output is around 5 L/min. (70 mL of blood × 70 beats/min). However, there is substantial variability: in highly perfused tissues (e.g., brain), typical circulation times are around 20 seconds, while for bone and joints it is up to 15 minutes. The average flow velocity in the aorta is about 20 cm/sec at rest, with a peak of about 1.5 m/sec. Under conditions of heavy exercise, these values are markedly increased.

The physical basis of diffusion is Brownian motion. Due to this phenomenon, more molecules in a homogeneous medium diffuse on the average from locations with a high concentration to locations where the concentration is lower than in the opposite direction. To obtain an approximate relation for typical diffusion times and distances, we outline a simple one-dimensional model that is based on the stochastic nature of diffusion (Joos and Freeman[158]). To this end, we start from an arbitrary distribution of a substance in a homogeneous solvent (Fig. 31). $N(x, t)$ denotes the concentration of the solute (number of particles per volume) at location x at time t. The particles execute a Brownian motion that is translationally invariant; that is, the probability that a particle moves from the location x to x' during the time interval ΔT due to Brownian motion depends only on the distance $d = x' - x$ and is independent of x itself, and, in addition, it is a function of ΔT. Let this probability that is constant in time be denoted by $p_{\Delta T}(x' - x)$. It is also symmetrical, $p_{\Delta T}(x' - x) = p_{\Delta T}(-[x' - x])$, because the Brownian motion acts equally in both directions.

At time $t' = t + \Delta T$ the number of particles at location x' is

$$N(x', t + \Delta T) = \int_{-\infty}^{\infty} N(x, t) \cdot p_{\Delta T}(x' - x)\, dx \qquad (106)$$

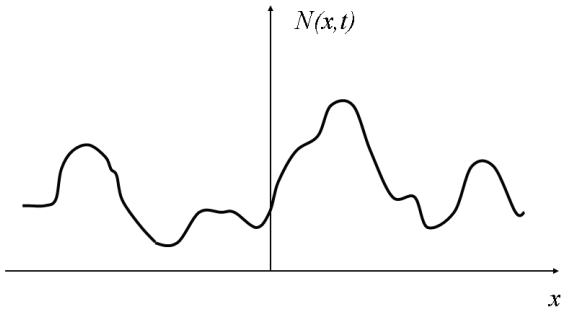


FIGURE 31. Arbitrary distribution of particles with density $N(x, t)$ in a homogeneous onedimensional diffusion model.

In this formulation it is assumed that the number of collisions associated with the Brownian motion during Δt is sufficiently large such that a local equilibrium always exists.

After the transformation $X = x' - x$ we arrive at

$$N(x', t + \Delta T) = \int_{-\infty}^{\infty} N(x' - X, t) \cdot p_{\Delta T}(X)\, dX \qquad (107)$$

This relation is now subjected to a Taylor expansion

$$\begin{aligned} & N(x', t) + \Delta T \left.\frac{\partial N}{\partial t}\right|_t + \ldots \\ &= \int_{-\infty}^{\infty} \left[N(x', t) - X \left.\frac{\partial N(x,t)}{\partial x}\right|_{x=x'} + \frac{1}{2} X^2 \left.\frac{\partial^2 N(x,t)}{\partial x^2}\right|_{x=x'} - \ldots \right] p_{\Delta T}(X)\, dX \end{aligned} \qquad (108)$$

Because of the symmetry property, $\int_{-\infty}^{\infty} X\, p_{\Delta T}(X) dX = 0$ such that the second term on the right hand side of Equation 108 disappears. Furthermore, because $p_{\Delta T}$ is a probability, $\int_{-\infty}^{\infty} p_{\Delta T}(X) dX = 1$. Upon discarding higher-order terms and retaining only the leading terms, Equation 108 reads

$$\frac{\partial N}{\partial t} = \frac{1}{2\Delta T} \frac{\partial^2 N}{\partial x^2} \int_{-\infty}^{\infty} X^2\, p_{\Delta T}(X) dX = \frac{1}{2\Delta T} \frac{\partial^2 N}{\partial x^2} \langle X^2 \rangle \qquad (109)$$

whereby $\langle X^2 \rangle$ denotes the squared average displacement which a particle undergoes during ΔT. Equation 109 is known as Fick's diffusion equation

$$\frac{\partial N}{\partial t} = D \frac{\partial^2 N}{\partial x^2} \qquad (110)$$

with the diffusion constant

$$D = \frac{\langle X^2 \rangle}{2\Delta T} \qquad (111)$$

If diffusion constants are known from measurements, then the typical time intervals necessary for diffusive transport over a distance *d* can be estimated using the relation shown in Equation 111. Table IV.1 has been derived from tabulated diffusion constants for free diffusion in water.

TABLE 1. Typical Diffusion Constants of Molecules in Water

D (m^2/sec)	$d = \sqrt{\langle X^2 \rangle}$ (μm)	$\Delta T^{\dagger}$ (sec)	Molecular Weight (D)
10^{-9} 10^{-9}	10 100	0.1 10	40
10^{-10} 10^{-10}	10 100	1 100	1000
10^{-9} 10^{-9}	10 100	10 1000	20'000

†Representative time intervals Δ*T*, which are necessary for molecules to diffuse on the average over a distance *d*, were estimated from Equation 111.

D, diffusion constants

It can be seen that for short distances diffusion may in fact be rapid. In biological tissues, diffusion distances are determined by the density of capillaries, which varies greatly from tissue to tissue. For well-perfused tissues, capillary-to-capillary distances are on the order of 20 μm to 100 μm; accordingly, diffusion is sufficiently effective to keep homeostasis. It should nevertheless be noted that large molecules diffuse in real tissues more slowly than in free water due to the presence of fibrous networks and other large molecules.

Different conditions are encountered in tissues that are weakly perfused, such as cartilage (no vascularization) or bone. Forced convection due to the action of muscles or deformation is an important additional transport mechanism in such tissues (likewise in venous and lymph flow). We consider compact bone in the following section.

B. Forced Convection

In typical compact bone, the bone matrix is essentially impermeable for fluids. Haversian (compact) bone consists of essentially cylindrically shaped, longitudinally arranged osteons. A single central capillary perfuses an osteon (diameter typically 200–400 μm). The bone cells (osteocytes) are distributed throughout the shell-like solid structure of the osteons and reside in openings within the bone matrix called lacunae (diameter a few micrometers). The lacunae, in turn, are connected by a dense network of canaliculi, through which the metabolic activity of the osteocytes is supported (Fig. 32). Because the canaliculi have diameters of less than about 400 nm, and osteon-osteon distances are typically around 30 μm, diffusion is by far insufficient to provide adequate transport capacity. This can be seen from a simple one-dimensional model (Fig. 33).

If we assume that there is only diffusive transport, we will find that the canaliculi will not lead to a sufficient physiologic transport volume. In order to demonstrate this, let, at time $t = 0$, a sample substance be injected at the entrance of the one-dimensional conduit (coordinate z) such that the concentration there is c_0, constant in time, in the otherwise substance-free channel. Accordingly, the boundary conditions for the concentration $c(z, t)$ to be determined are

$$\begin{aligned} &t<0: \quad c(z,t) = 0 \\ &t>0: \quad c(z,t) = c_0 \quad (constant) \\ &t>0: \quad c(z=\infty,t) = 0 \end{aligned} \tag{112}$$

In order to solve the diffusion equation,

$$\frac{\partial^2 c}{\partial z^2} = \frac{1}{D}\frac{\partial c}{\partial t} \tag{113}$$

It is subjected to a Laplace transform

$$\frac{\partial^2 \tilde{c}}{\partial z^2} = \frac{s}{D}\tilde{c} \tag{114}$$

Thereby, s denotes the variable associated with the Laplace transform, and $\tilde{c}$ the Laplace transform of c. The transformed boundary conditions read

$$\begin{aligned} \tilde{c}(s, z=0) &= \frac{c_0}{s} \\ \tilde{c}(s, z=\infty) &= 0 \end{aligned} \tag{115}$$

Upon solving Equation 114 and performing the inverse transform, one obtains

$$c(z,t) = c_0 \, erfc\left(\frac{z}{2\sqrt{Dt}}\right) \tag{116}$$

($erfc = 1 - erf$). The flow per unit area, disregarding wall effects, $Q = -D\frac{dc}{dz}$ can be calculated from

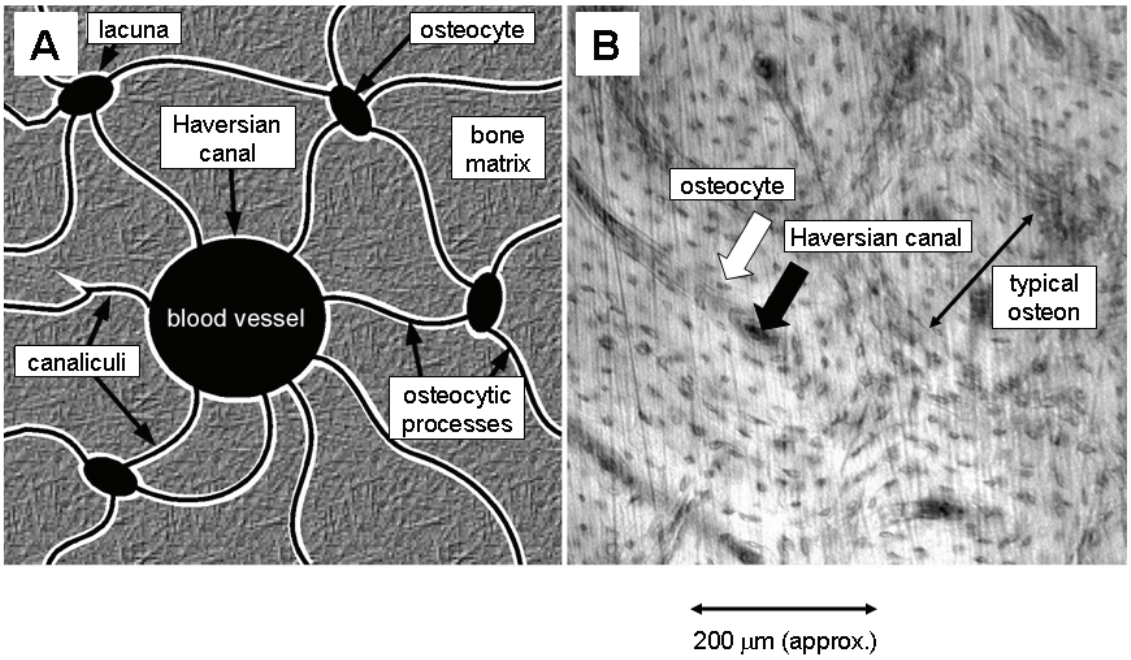


FIGURE 32. A, Schematic view of compact (Haversian) bone highlighting the lacunocanalicular system. This type of bone consists of longitudinally arranged osteons (typical diameter 200 μm), which have a cylindrical shell-like structure around a central blood vessel. The osteocytes are housed in the lacunae and are connected by numerous canaliculi. Metabolic traffic (nutrients, waste products) associated with osteocytic activity occurs essentially through the canalicular system. B, Microscopic view. (Adapted from Knapp HF, Reilly GC, Stemmer A, Niederer P, Knothe Tate ML. Development of preparation methods for and insights obtained from atomic force microscopy of fluid spaces in cortical bone. Scanning. 2002;24:25–33.)

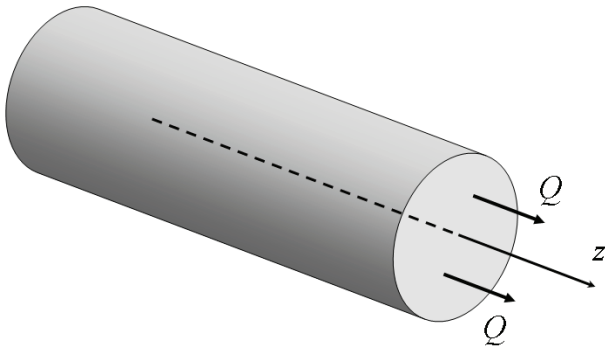


FIGURE 33. Diffusion along a cylindrical conduit.

Equation 116 as

$$Q = c_0 \left(\frac{D}{\pi t}\right) e^{-\left(\frac{z}{2\sqrt{Dt}}\right)^2} \qquad (117)$$

using the differentiation properties of the error function.

The exponential factor for $z = 10\,\mu m$, $D = 10^{-8}\,cm^2/sec$, $t = 1\,sec$, say, is close to zero. It has in fact been shown that the perfusion of compact bone is mainly due to forced convection: as a result of the loading patterns to which bones are exposed during normal daily activity (walking, working, etc.) the bone matrix undergoes small deformations. Because the fluid contained in the lacuno-canalicular system is nearly incompressible, load-induced pressure gradients build up, which set the fluids in motion (Fig. 34). Load-induced fluid flow or forced convection has been found to be an extremely powerful and effective transport mechanism.

There are other examples of forced convection. For example, when we get up in the morning and stretch our limbs, we build up pressure gradients in the various tissues, which stimulates interstitial perfusion and lymph flow. Internal contractile forces enhance transport likewise.

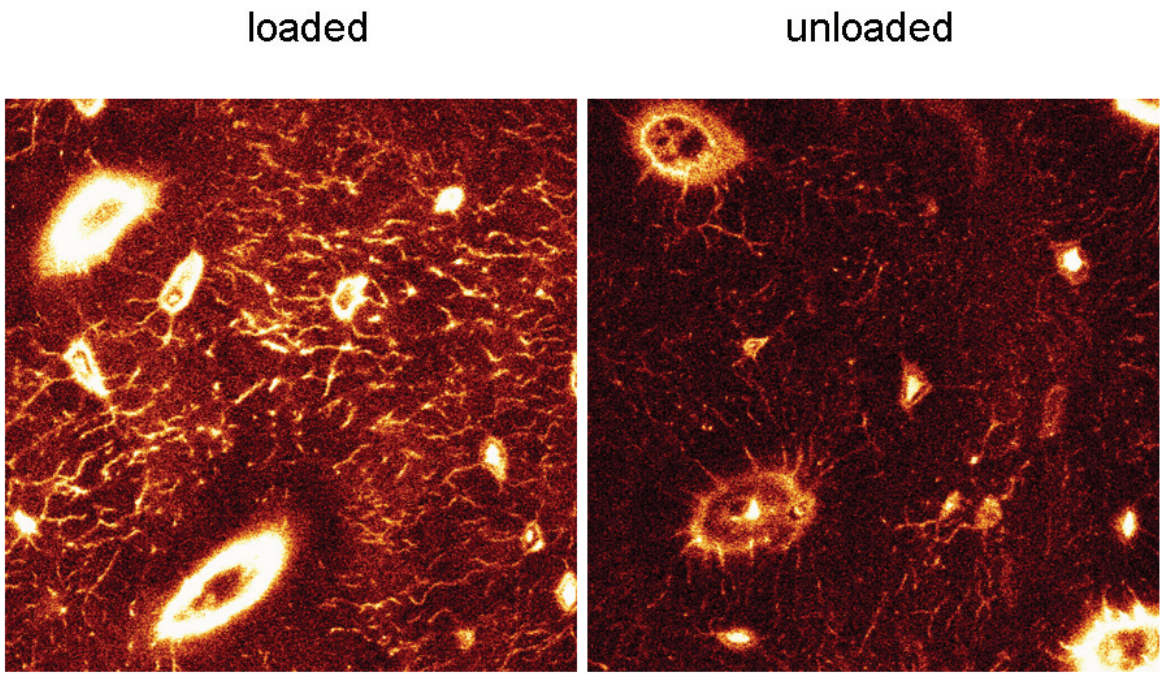


FIGURE 34. A, Forced convection in a rat tibia, in vivo experiment. The anesthetized animal was systemically perfused for 5 minutes with a Procion red solution (inert fluorescent dye, molecular weight 200 D) prior to the loading experiment. While one of the bones remained unloaded (right), the contralateral side (left) was subjected to a number of load cycles that correspond to typical walking conditions. The increased amount of Procion red manifests itself (intensity of fluorescence) on the loaded side and documents forced convection (magnification 1000×).

C. Osmosis

Osmosis is a particularly important phenomenon in biology. It is associated with diffusion and occurs in systems that are subdivided into fluid compartments separated by semipermeable membranes; that is, barriers that are permeable selectively for certain molecules only. While the solvent, in biology mostly water, can diffuse freely between the compartments through the diffusive barrier, the solute cannot, because the molecules are too big to penetrate the membrane. Whenever concentration differences exist across the barrier, a pressure difference develops in thermodynamic equilibrium called osmotic pressure. This pressure is an important driving mechanism in many biological transport processes.

In what follows, two fluid spaces are considered that are separated by a semipermeable membrane. The two spaces contain mixtures with different concentrations, whereby the membrane is permeable for the solvent and impermeable for the solute. We perform an analysis based on classical thermodynamics that will lead us to van't Hoff's law. In the derivation presented here, it is valid for diluted solutions and small molecules (more general treatments can be found in Atkins and de Paula[167] or Landau and Lifshitz.[168]

We start by considering a homogeneous system consisting of one substance only. According to the second law of thermodynamics, the change of entropy, dS, of such a system when subjected to an infinitesimal state change, can be written as

$$dS \geq \sum_k X_k \, dx_k \tag{119}$$

X_k thereby denote "intensive" quantities (independent of the amount of material such as the temperature, T, or the pressure, p), while the associated x_k are "extensive" quantities (dependent on the amount of material such as the internal energy, U, or the volume, V). The equal sign holds for reversible processes, whereas in case of irreversible state changes, the entropy increases.

If the system can fully be described by the aforementioned four quantities, i.e., temperature, pressure, internal energy and volume, Equation 118 reads

$$dS \geq \frac{1}{T} dU + \frac{p}{T} dV \tag{119}$$

In this formulation, the extensive quantities U, V act as independent variables. For many biological processes, in particular transport processes, it is more useful to have a representation with the intensive quantities as independent variables. This is reached by the following transformation ("contact" transformation), which introduces the Gibbs potential, G,

$$G = U + pV - TS \tag{120}$$

Upon calculating the total differential, dG, and using Equation 119, one obtains

$$dG(p, T) \leq -S\,dT + V\,dp \tag{121}$$

The new quantity, the Gibbs potential G, has the desired form. Among others, from Equation 121 follows

$$\left.\frac{\partial G}{\partial p}\right|_T = V \tag{122}$$

a relation that will be of use later.

After these introductory remarks, we turn our attention now to homogeneous mixtures of various substances, i ($i = 1,..., k$) in the same solvent. The number of moles of each substance in the mixture, n_i, serve as additional extensive quantities. The associated intensive quantities, μ_i, are denoted as chemical potential of each substance. Following Equation 118, the natural extension of Equation 122 reads

$$dG(p, T, n_i) \leq -S\,dT + V\,dp + \sum_k \mu_i\, n_i \tag{123}$$

It is furthermore useful to relate all intensive quantities for each substance to the number of moles, that is

$$v_i = \left.\frac{\partial V}{\partial n_i}\right|_{p,T,n_j}, \quad s_i = \left.\frac{\partial S}{\partial n_i}\right|_{p,T,n_j} \quad (j \neq i) \tag{124}$$

such that Equation 123 can be written as

$$dG \leq \sum_k \left[(-s_i\,dT + v_i\,dp)n_i + \mu_i\,dn_i\right] \tag{125}$$

From Equation 125 it follows that the chemical potential, μ_i, has the meaning of the molar Gibbs potential, that is,

$$\mu_i = \left.\frac{\partial G}{\partial n_i}\right|_{p,T,n_j}, \text{ and } G = \sum_k \mu_i\, n_i \tag{126}$$

(Compare Equation 125 with $dG = \sum_k \left[d\mu_i\, n_i + \mu_i\, dn_i\right]$ from Equation 126).

As mentioned before, the derivation presented here is restricted to low concentrations, which implies that the molecules of the various solutes are sufficiently far away from each other such that eventual mutual interactions can be neglected. Accordingly, for the solutes, use will be made of the pressure dependence of the chemical potential of an ideal gas. From

$$\left.\frac{\partial \mu_i}{\partial p}\right|_T = v_i \text{ (see Equations 125 and 126), one obtains} \tag{127}$$

$$\mu_i(T, p_i) - \mu_i^0(T)\Big|_{p_0} = \left.\int_{p_0}^{p_i} v_i\,dp\right|_{T=const} = \int_{p_0}^{p_i} \frac{RT}{p}\,dp = RT\log\left(\frac{p_i}{p_0}\right) \tag{128}$$

where the ideal gas law, $p_i\,v_i = RT$, has been applied. p_0 denotes some arbitrary reference state, while p_i is the partial pressure of the component i in the mixture. According to Dalton's law, the relations

$$p = \sum_k p_i \quad \text{and} \quad \frac{p_i}{p} = \frac{n_i}{n} \quad \left(n = \sum_k n_i\right) \tag{129}$$

hold. In this equation, the solvent is included. Because this is a liquid, its chemical potential can not directly be derived from the ideal gas equation.

However, a useful relation can be obtained from the equilibrium between the liquid and the associated vapor phase. The latter can again be approximated by an ideal gas because the molecules are distributed sparsely in air. When dn_{li} moles of the liquid phase evaporate into the vapor phase, the Gibbs potential changes according to

$$dG_{li} = (\mu_v - \mu_{li})dn_{li} \tag{130}$$

where μ_v denotes the chemical potential of the vapor and μ_{li} the chemical potential of the liquid phase, respectively. The Gibbs potential changes spontaneously in an irreversible fashion until thermodynamic equilibrium is reached, which is the case when $dG_{li} = 0$. The chemical potentials are therefore equal,

that is,

$$\mu_{li} = \mu_{li}\big|_{p_0} + RT \log\left(\frac{p_{li}}{p_0}\right) \tag{131}$$

if it is assumed, as indicated above, that the vapor obeys the ideal gas relation.

In the mixture, the total number of moles is n, and the total pressure is p. From this, n_{li} and p_{li}, respectively, are due to the (liquid) solvent. The total pressure, p, is now introduced by

$$\begin{aligned}\mu_{li} &= \mu_{li}\big|_{p_0} + RT\left(\log p_{li} - \log p_0 + \log p - \log p\right) \\ &= \mu_{li}\big|_{p_0} + RT\left[\log\left(\frac{p_{li}}{p}\right) + RT\left(\frac{p}{p_0}\right)\right] = \mu_{li}\big|_{p} + RT\log\left(\frac{p_{li}}{p}\right)\end{aligned} \tag{132}$$

Utilizing Dalton's law (Equation 129), finally, the relation holds

$$\mu_{li} = \mu_{li}\big|_{n_{li}=n} + RT\log\left(\frac{n_{li}}{n}\right) \tag{133}$$

Equation 133 allows us now to derive van't Hoff's law. To this end, we consider the system shown in Figure 35. It consists of a **U**-shaped tube separated into two halves by a semipermeable membrane at the bottom. The tube is filled with a diluted solution and the membrane is permeable for the solvent but impermeable for the solute. The one side (side 1) of the tube contains a solution consisting of n^1_{so} moles of solute in n^1_{li} moles of solvent $(n^1 = n^1_{so} + n^1_{li})$, while on the other (side 2) there is only solvent. It is empirically observed that the pressure on side 1 is higher in comparison with side 2 by an amount, π, called osmotic pressure.

Because the solvent can move freely through the membrane, the chemical potentials of the solvent on both sides, μ^1_{li}, μ^2_{li}, respectively, are the same, once an equilibrium is established. Otherwise, fluid

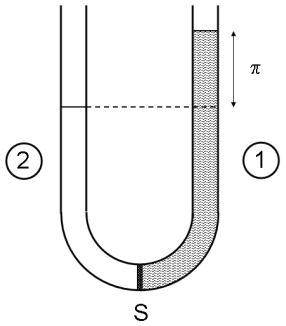


FIGURE 35. Osmotic pressure π. S, semipermeable membrane.

is exchanged through the membrane until the total Gibbs potential reaches a minimum. Therefore,

$$\mu^2_{li\,(\text{pure solvent})}(p) = \mu^1_{li\,(\text{solution})}\left(\frac{n^1_{li}}{n}, p + \pi\right) \tag{134}$$

The right side of Equation 134 can be evaluated according to Equations 128 and 133:

$$\begin{aligned}\mu^2_{li}(p) &= \mu^1_{li}(p+\pi) + RT\log\left(\frac{n^1_{li}}{n}\right) \\ &= \mu^1_{li}(p) + \int_p^{p+\pi} v_{li}\,dp + RT\log\left(\frac{n^1_{li}}{n}\right)\end{aligned} \tag{135}$$

The molar volume, v_{li}, of a liquid is approximately constant under the pressures of interest here and, because the solution is diluted, $v_{li} \approx v$, such that the integral in Equation 135 equals $v\,\pi$. Furthermore, $\mu^1_{li}(p)$ and $\mu^2_{li}(p)$ are the same, One arrives therefore at

$$v\pi = -RT\log\left(\frac{n^1_{li}}{n}\right) = -RT\log\left(1 - \frac{n^1_{so}}{n}\right) \approx RT\left(\frac{n^1_{so}}{n}\right) \tag{136}$$

where only the first term of the Taylor expansion has been retained. Because total volume $V = vn$,

$$\pi = RT\left(\frac{n^1_{so}}{V}\right) = RT\left(\frac{c_{so}}{M_{so}}\right) \text{(van't Hoff's law)} \tag{137}$$

whereby, with M_{so} denoting the molecular weight of the solutes, and c_{so} the concentration, the relation has been used,

$$c = \left(\frac{n^1_{so}\,M_{so}}{V}\right) \tag{138}$$

Van't Hoff's law (Equation 137) holds for diluted solutions and solutes with small molecular weights. For increasing M_{so}, π is empirically found to decrease. The reason is that the ideal gas approximation is not valid in the case of large molecules, because in such substances the intermolecular potentials cannot be disregarded (in an ideal gas, there are no interactions between the particles other than collisions causing equilibrium; there are in particular no intermolecular potentials).

Osmosis still occurs in the case of solutions involving large molecules (proteins), but its effect cannot be derived in a straightforward manner. A method called virial expansion, which is based on a correlation analysis (second-order correlations relate

to pair-wise interaction of molecules, the third-order to triple interactions, etc.) allows an approximation of the osmotic or oncotic pressure (sometimes called "oncotic" in order to distinguish it from the previous form) according to

$$\pi = RT\left(\frac{c}{M_s} + A_2 c^2 + A_3 c^3 + \ldots\right) \quad (139)$$

The A_i are called virial coefficients.

In the case of blood, in which water is the solvent, the osmotic pressure caused by the electrolytes in solution (Na^+, Ca^{++}, $C1^-$, etc., in total ~1 g/100 mL of plasma) is about 750 kPa, which is in agreement with van't Hoff's law. However, this enormous pressure does not manifest itself in the human circulation, because the capillaries are generally permeable for these (small) ions. In the larger blood vessels, osmotic pressures do not develop across the vessel wall because it is essentially impermeable. An important exception consists of the vasculature and microcirculation of the central nervous system, where the capillaries are impermeable as a rule (blood-brain barrier). Exchange of substances across this barrier is essentially based on active transport processes.

In addition to the small ions mentioned above, the blood contains also large molecules, the plasma proteins (molecular weight from ~10,000–300,000 D). As a rule of thumb, "small" and "large" refers to a molecular weight of less than or more than 5000 D, respectively. Capillary walls, in particular on the arterial side, including the endothelial glycocalix are to a great extent impermeable for large molecules such as proteins. Accordingly, an oncotic pressure develops, which for a plasma protein concentration of about 7.5 g/100 mL of plasma is around 3.3 kPa. This is the basis of Starling's hypothesis, which serves as a model for the description of capillary exchange in tissues other than the central nervous system.

Starling's formula for the fluid flow through the capillary wall, *J*, is usually given as

$$J = p_c\left(\Delta p - \varepsilon \Delta\pi\right) \quad (140)$$

whereby p_c is a parameter characterizing the permeability of the wall, Δp the hydrostatic pressure difference (difference between the fluid pressure inside the capillary and the interstitial space), and $\Delta\pi$ the oncotic pressure difference (difference due to the different plasma protein concentration in the intra- and extravascular space). ε is a correction factor ("reflection coefficient") that takes into account the variable, albeit usually small, leakage of large proteins through the capillary wall and the fact that not all proteins are effective in retaining water.

Starling's hypothesis has been substantiated qualitatively in many tissues, mostly in animal experiments.[169-173] Nevertheless, the microcirculation is different from tissue to tissue and depends largely on the physiological and biochemical characteristics. In addition to the passive fluid exchange approximated by Equation 140, there are numerous active transport processes involving carrier molecules that are tissue specific and modulate capillary exchange and extravascular (interstitial) transport.

REFERENCES

1. Moon P, Spencer DE. Field theory handbook. Berlin: Springer; 1988.
2. Courant R, Hilbert D. Methods of mathematical physics. Vol 1/2. New York: Wiley-Interscience; 2007/2008.
3. Bronson R. Schaum's Outline of differential equations. 3rd ed. New York: McGraw Hill; 2009.
4. Gallo LM, Fushima K, Palla S. Mandibular helical axis pathways during mastication. J Dent Res. 2000;79:1566–72.
5. Woltring HJ, Long K, Osterbauer PJ, Fuhr AW. Instantaneous helical axis estimation from 3-D video data in neck kinematics for whiplash diagnostics. J Biomech. 1994;27:1415–25.
6. Kane TR, Levinson DA. Multibody dynamics. ASME Trans J Appl Mech. 1983;50:1071–8.
7. Kane TR, Likins PW, Levinson DA. Spacecraft dynamics. New York: McGraw-Hill; 1983.
8. Wittenburg J. Dynamics of multibody systems. Berlin: Springer; 2008.
9. Giansanti D, Maccioni G, Macellari V. The development and test of a device for the reconstruction of 3-D position and orientation by means of a kinematic sensor assembly with rate gyroscopes and accelerometers. IEEE Trans Biomed Eng. 2005;52:1271–7.
10. Takeda R, Tadano S, Natorigawa A, Todoh M, Yoshinari S. Gait posture estimation by wearable acceleration and gyro sensor. In: Dössel O, Schlegel WC, editors. IFMBE Proceedings, Vol. 25/3. World Congress on Medical Physics and Biomedical Engineering. Berlin: Springer; 2009. p. 111–4.
11. Zijlstra W, Bisseling R. Estimation of hip abduction moment based on body fixed sensors. Clin Biomech. 2004;19:819–27.

12. Biryukova EV, Roby-Brami A, Frolov AA, Mokhtari M. Kinematics of human arm reconstructed from spatial tracking system recordings. J Biomech. 2000;33:985–95.
13. Fan J, El-Kwae EA, Hacid MS, Liang F. Novel tracking-based moving object extraction algorithm. J Electron Imaging. 2002;11:393.
14. Green RD, Guan L, Burne JA. Video analysis of gait for diagnosing movement disorders. J Electron Imaging. 2000;9:16.
15. Pascolo P. Computation of rigid body motion parameters from video-based measurements. In: Tarantino U, Perugia D, Campanacci G, Pennestri E, editors. Biomechanics and sports. CISM Courses and Lectures. 2004;473:11–25.
16. Spoor CW, Veldpaus FE. Rigid body motion calculated from spatial co-ordinates of markers. J Biomech. 1980;13:391–3.
17. Pearle AD, Kendoff D, Musahl V. Perspectives on computer-assisted orthopaedic surgery: movement toward quantitative orthopaedic surgery. J Bone Joint Surg. 2009;91A:7–12.
18. Nottmeier EW, Seemer W, Young PM. Placement of thoracolumbar pedicle screws using three-dimensional image guidance: experience in a large patient cohort clinical article. J Neurosurg-Spine. 2009;10:33–9.
19. Haaker RG, Stockheim M, Kamp M, Proff G, Breitenfelder J, Ottersbach A. Computer-assisted navigation increases precision of component placement in total knee arthroplasty. Clin Orthop Rel Res. 2005;433:152–9.
20. Tresadern PA, Reid ID. Camera calibration from human motion. Image Vis Comput. 2008;26:851–62.
21. Aggarwal JK, Cai Q. Human motion analysis: A review. Comp Vision and Image Understanding. 1999;73:428–40.
22. Ambrosio JAC, editor. Advances in computational multibody systems. Dordrecht: Springer; 2005.
23. Bogert A, Schamhardt HC, Crowe A. Simulation of quadrupedal locomotion using a rigid body model. J Biomech. 1989;22:33–41.
24. Huang SC. Analysis of human body dynamics in simulated rear-end impacts. Hum Mov Sci. 1998;17:821–38.
25. Okita N, Meyers SA, Challis JH, Sharkey NA. An objective evaluation of a segmented foot model. Gait & Posture. 2009;30:27–34.
26. Pandy MG. Computer modeling and simulation of human movement. Ann Rev Biomed Eng. 2001;3:245–73.
27. Wojtra M. Multibody simulation model of human walking. Mech Struct Machines. 2003;3:357–79.
28. Hase K. Computational rehabilitation engineering with a walking model. In: Kajzer J, Tanaka E, Yamada H, editors. Hum Biomech & Injury Prev. Berlin; Springer: 2000, p. 217–22.
29. Untaroiu CD, Crandall JR, Takahashi Y, Okamoto M, Ito O, Fredriksson R. Analysis of running child pedestriansimpacted by a vehicle using rigid-body models and optimization techniques. Safety Sci. 2010;48:259–76.
30. Andersen MS, Damsgaard M, Rasmussen J. Kinematic analysis of over-determinate biomechanical systems. Comp Meth in Biomech Biomed Eng. 2009;12:371–84.
31. Callaghan JP, Patla AE, McGill SM. Low back three-dimensional joint forces, kinematics, and kinetics during walking. Clin Biomech. 1999;14:203–16.
32. Chen G. Induced acceleration contributions to locomotion dynamics are not physically well defined. Gait & Posture. 2006;23:37–44.
33. Van den Bogert AJ, Read L, Nigg BM. A method for inverse dynamic analysis using accelerometry. J Biomech. 1996;29:949–54.
34. Landau LD, Lifshitz EM. Course in theoretical physics, Vol. 3. Quantum mechanics. Oxford: Butterworth-Heinemann; 2003.
35. Peleg Y, Pnini R, Zaarur E, Hecht E. Schaum's outline of quantum mechanics. 2nd ed. New York: McGraw Hill; 2010.
36. Phillips AC. Introduction to quantum mechanics. Chichester, England: J. Wiley & Sons Ltd.; 2003.
37. Frenkel D, Smit B. Understanding molecular simulation: from algorithms to applications. London: Academic Press; 2002.
38. Jensen F. Introduction to computational chemistry. 2nd ed. Chichester, England: J. Wiley & Sons Ltd.; 2006.
39. Leach AR. Molecular modelling: principles and applications. 2nd ed. Princeton, NJ: Prentice Hall; 2001.
40. Balestrini JL, Skorinko JK, Hera A, Gaudette GR, Billiar KL. Applying controlled non-uniform deformation for in vitro studies of cell mechanobiology. Biomech Model Mechanobiol. 2010;9:329–44.
41. Bao G, Kamm RD, Thomas W, Hwang W, Fletcher DA, Grodzinsky AJ, Zhu C, Mofrad MRK. Molecular biomechanics: The molecular basis of how forces regulate cellular function. Cell Mol Bioeng. 2010;3:91–105.
42. Chen CS, Tan J, Tien J. Mechanotransduction at cell-matrix and cell-cell contacts. Ann Rev Biomed Eng. 2004;6:275–302.
43. Chen CS. Mechanotransduction--a field pulling together? J Cell Sci. 2008;121:3285–92.
44. Dahl KN, Kalinowski A, Pekkan K. Mechanobiology and the microcirculation: cellular, nuclear and fluid mechanics. Microcirculation. 2010;17:179–91.
45. Hamill OP, Martinac B. Molecular basis of mechanotransduction in living cells. Physiol Rev. 2001;81:685–740.
46. Huang H, Kamm RD, Lee RT. Cell mechanics and mechanotransduction: Pathways, probes and physiology. Am J Physiol Cell Physiol. 2003;287:C1–11.
47. Humphrey JD. Vascular adaptation and mechanical homeostasis at tissue, cellular, and sub-cellular levels. Cell Biochem Biophys. 2008;50:53–78.
48. Humphrey JD. Vascular mechanics, mechanobiology and remodeling. J Mech Med Biol. 2009;9:243–57.
49. Ingber DE. Cellular mechanotransduction: Putting all the pieces together again. FASEB J. 2006;20:811–27.
50. Ingber DE. Mechanobiology and diseases of mechanotransduction. Ann Med. 2003;35:564–77.
51. Katsumi A, Orr AW, Tzima E, Swartz MA. Integrins in mechanotransduction. J Biol Chem. 2004;279:12001–4.

52. Mofrad MRK, Kamm RD, editors. Cellular mechanotransduction. diverse perspectives from molecules to tissues. Cambridge: Cambridge University Press; 2009.
53. Pedersen JA, Lichter S, Swartz MA. Cells in 3D matrices under interstitial flow: Effects of extracellular matrix alignment on cell shear stress and drag forces. J Biomech. 2010;43:900–5.
54. Silver FH, Siperko LM, Seehra GP. Mechanobiology of force transduction in dermal tissue. Skin Res Technol. 2003;9:3–23.
55. Socci L, Pennati G, Gastaldi D, Vena P. Modeling and mechanobiology of cerebral aneurysms. J Appl Biomat Biomech. 2008;6:63–71.
56. Suresh S. Biomechanics and biophysics of cancer cells. Acta Biomaterials. 2007;3:413–38.
57. Tarbell JM, Pahakis MY. Mechanotransduction and the glycocalix. J Int Med. 2006;259:339–50.
58. Ulrich TA, Jain A, Tanner K, MacKay JL, Kumar S. Probing cellular mechanobiology in three-dimensional culture with collagen-agarose matrices. Biomaterials. 2010;31:1875–84.
59. Wang JH, Thampatty BP. An introductory review of cell mechanobiology. Biomech Model Mechanobiol. 2006;5:1–16.
60. Wang N, Butler JP, Ingber DE. Mechanotransduction across the cell surface and through the cytoskeleton. Science. 1993;260:1124–7.
61. Wang YX, Shyy JYJ, Chien S. Fluorescence proteins, live-cell imaging, and mechanobiology: Seeing is believing. Ann Rev Biomed Eng. 2008;10:1–38.
62. Wozniak MA, Chen CS. Mechanotransduction in development: a growing role for contractility. Nature Rev Mol Cell Biol. 2009;10:34–43.
63. Fritton SP, Weinbaum S. Fluid and solute transport in bone: flow-induced mechanotransduction. Ann Rev Fluid Mech. 2009;41:347–74.
64. Han YF, Cowin SC, Schaffler MB, Weinbaum S. Mechanotransduction and strain amplification in osteocyte cell processes. Proc Nat Acad Sci U S A. 2004;101:16689–94.
65. Isaksson H, Grongroft I, Wilson W, van Donkelaar CC, van Rietbergen B, Tami A, Huiskes R, Ito K. Remodeling of fracture callus in mice is consistent with mechanical loading and bone remodeling theory. J Orthopaedic Res. 2009;27:664–72.
66. Jacobs CR. The mechanobiology of cancellous bone structural adaptation. J Rehab Res Devel. 2000;37:209–16.
67. Kim BG, Kwag JH, Kim HS, Kim KH, Kim CH. Bone cell response to fluid shear stress and cyclic compressive strain in 3D trabecular bone. Tissue Eng Regenerative Med. 2009;6:540–8.
68. Mullender M, El Haj AJ, Yang Y, van Duin MA, Burger EH, Klein-Nulend J. Mechanotransduction of bone cells in vitro: mechanobiology of bone tissue. Med Bio Eng Comput. 2004;42:14–21.
69. Mulvihill BM, Prendergast PJ. Mechanobiological regulation of the remodelling cycle in trabecular bone and possible biomechanical pathways for osteoporosis. Clin Biomech. 2010;25:491–8.
70. Papachroni KK, Karatzas DN, Papavassiliou KA, Basdra EK, Papavassiliou AG. Mechanotransduction in osteoblast regulation and bone disease. Trends Mol Med. 2009;15:208–16.
71. Qi MC, Zou SJ, Han LC, Zhou HX, Hu J. Expression of bone-related genes in bone marrow MSCs after cyclic mechanical strain: Implications for distraction osteogenesis. Int J Oral Sci. 2009;1:143–50.
72. Riddle RC, Donahue HJ. From streaming potentials to shear stress: 25 years of bone cell mechanotransduction. J Orthop Res. 2009;27:143–9.
73. Turner CH, Warden SJ, Bellido T, Plotkin LI, Kumar N, Jasiuk I, Danzig J, Robling AG. Mechanobiology of the skeleton. Sci Signaling. 2009;2(68):pt3.
74. Van der Meulen MCH, Moro M, Kiratli BJ, Marcus R, Bachrach RK. Mechanobiology of femoral neck structure during adolescence. J Rehab Res Devel. 2000;37:201–8.
75. You LD, Temiyasathit S, Lee PL, Kim CH, Tummala P, Yao W, Kingery W, Malone AM, Kwon RY, Jacobs CR. Osteocytes as mechanosensors in the inhibition of bone resorption due to mechanical loading. Bone. 2008;42:172–9.
76. Beaupre GS, Stevens SS, Carter DR. Mechanobiology in the development, maintenance, and degeneration of articular cartilage. J Rehab Res Devel. 2000;37:145–51.
77. Carter DR, Wong M. Modelling cartilage mechanobiology. Phil Trans Roy Soc London Ser B Biol Sci. 2003;358:1461–71.
78. Kelly DJ, Jacobs CR. The role of mechanical signals in regulating chondrogenesis and osteogenesis of mesenchymal stem cells. Birth Defects Res Part C Embryo Today Revs. 2010;1:75–85.
79. Muir H. The condrocyte, architect of cartilage--biomechanics, structure, function and molecular-biology of cartilage matrix macromolecules. Bioessays. 1995;17:1039–48.
80. Stoltz JF, editor. Mechanobiology: cartilage and chondrocytes. Vol. 3. Amsterdam: IOS Press; 2004.
81. Nagel T, Kelly DJ. Mechano-regulation of mesenchymal stem cell differentiation and collagen organisation during skeletal tissue repair. Biomech Model Mechanobiol. 2010;9:359–72.
82. Chien S. Mechanotransduction and endothelial cell homeostasis: the wisdom of the cell. Am J Physiol Heart Circ Physiol. 2007;292:H1209–24.
83. Davies PF. Flow-mediated endothelial mechanotransduction. Physiol Rev. 1995;75:519–60.
84. Chiquet M, Gelman L, Lutz R, Maier S. From mechanotransduction to extracellular matrix gene expression in fibroblasts. Biochim Biophys Acta Mol Cell Res. 2009;1793:911–20.
85. Grillet N, Kazmierczak P, Xiong W, Schwander M, Reynolds A, Sakaguchi H, Tokita J, Kachar B, Muller U. The mechanotransduction machinery of hair cells. Sci Signal. 2009;2(85):pt5.

86. Kearney EM, Prendergast PJ, Campbell VA. Mechanisms of strain-mediated mesenchymal stem cell apoptosis. J Biomech Eng. 2008 Dec;130(6):061004.
87. Wang JHC, Thampatty BP. Mechanobiology of adult and stem cells. Int Rev Cell and Mol Biol. 2008;271:301–46.
88. Danos N, Staab KL. Can mechanical forces be responsible for novel bone development and evolution in fishes? J Appl Ichthyol. 2010;26:156–61.
89. Cao Y, Li HB. Engineered elastomeric proteins with dual elasticity can be controlled by a molecular regulator. Nature Nanotechnol. 2008;3:512–6.
90. James HM, Guth E. Statistical Thermodynamics of rubber elasticity. J Chem Phys. 1953;21:1039–49.
91. Von Castelmur E, Marino M, Svergunt DI, Kreplak L, Ucurum-Fotiadis Z, Konarev PV, Urzhumtsev A, Labeit D, Labeit S, Mayans O. A regular pattern of Ig super-motifs defines segmental flexibility as the elastic mechanism of the titin chain. Proc Nat Acad Sci U S A. 2008;105:1186–91.
92. Mase GE. Schaum's outline of continuum mechanics. New York: John Wiley & Sons; 1988.
93. Narasimhan MNL. Principles of continuum mechanics. New York: John Wiley & Sons; 1993.
94. Huang CY, Mow VC, Ateshian GA. The role of flow-independent viscoelasticity in the biphasic tensile and compressive responses of articular cartilage. J Biomech Eng Trans ASME. 2001;123:410–7.
95. Huang CY, Soltz MA, Kopacz M, Mow VC, Ateshian GA. Experimental verification of the roles of intrinsic matrix viscoelasticity and tension-compression nonlinearity in the biphasic response of cartilage. J Biomech Eng Trans ASME. 2003;125:84–93.
96. Holzapfel GA, Ogden RW, editors. Mechanics of biological tissues. Berlin: Springer; 2006.
97. YamadaH. Strength of biological materials. Philadelphia: Williams & Wilkins; 1970.
98. Dorri F, Niederer P, Lunkenheimer PP. A finite element model of the left ventricular systole. Comp Meth Biomech Biomed Eng. 2006;9:319–41.
99. Fung YC. Biomechanics: mechanical properties of living tissues. New York: Springer; 1993.
100. Cowin SC, editor. Bone mechanics handbook. Boca Raton, FL: CRC Press; 2001.
101. Holzapfel GA, Humphrey JD, editors. Biomechanics and modeling in mechanobiology. Berlin: Springer; 2006.
102. Holzapfel GA, Ogden RW. Constitutive modelling of arteries. Proc Roy Soc A Math Phys Eng Sci. 2010;466:1551–96.
103. Pioletti DP, Rakotomanana LR. Non-linear viscoelastic laws for soft biological tissues. Eur J Mech A Solids. 2000;19:749–59.
104. Yamashita J, Li X, Furman BR, Rawls HR, Wang X, Agrawal CM. Collagen and bone viscoelasticity: a dynamic mechanical analysis. J Biomed Mater Res. 2002;63:31–6.
105. Bausch AR, Ziemann F, Boulbitch AA, Jacobson K, Sackmann E. Local measurements of viscoelastic parameters of adherent cell surfaces by magnetic bead microrheometry. Biophys J. 1998;75:2038–49.
106. Braddock M, Schwachtgen JL, Houston P, Dickson MC, Lee MJ, Campbell CJ. Fluid shear stress modulation of gene expression in endothelial cells. News Physiol Sci. 1998;13:241–6.
107. Ambrosi D, Mollica F. On the mechanics of a growing tumor. Int J Eng Sci. 2002;40:1297–1316
108. Humphrey JD. Continuum biomechanics of soft biological tissues. Proc Roy Soc A. 2003;459:3–46.
109. Lim CT, Zhou EH, Quek ST. Mechanical models for living cells--a review. J Biomech. 2006;39:195–216.
110. Schmitt K-U, Niederer P, Muser M, Walz F. Trauma biomechanics. 3rd ed. Berlin: Springer; 2010.
111. Dumont ER. Bone density and the lightweight skeletons of birds. Proc Roy Soc B Biol Sci. 2010;277:2193–8.
112. Langerhans RB. Using biomechanics and ecology to predict the evolution of body form and swimming performance in fishes. Comp Biochem Physiol A Mol Integr Physiol. 2007;148:S53.
113. Sfakiotakis M, Lane DM, Davies JB, Review of fish swimming modes for aquatic locomotion. IEEE J Oceanic Eng. 1999;24:237–52.
114. Shadwick RE, Lauder GV, editors. Fish biomechanics. London: Academic Press; 2006.
115. Wakeling JM. Biomechanics of fast-start swimming in fish. Comp Biochem Physiol A Mol Integr Physiol. 2001;131:31–40.
116. Altshuler DL, Dudley R. The physiology and biomechanics of avian flight at high altitude. Integr Comp Biol. 2006;46:62–71.
117. Faruque I, Humbert JS. Dipteran insect flight dynamics. Part 1. Longitudinal motion about hover. J Theor Biol. 2010;264:538–52.
118. Tobalske BW, Biewener AA, Warrick DR, Hedrick TL, Powers DR. Effects of flight speed upon muscle activity in hummingbirds. J Exp Biol. 2010;213:2515–23.
119. Tobalske BW. Biomechanics of bird flight. J Exp Biol. 2007;210:3135–46.
120. Almeras T, Fournier M. Biomechanical design and long-term stability of trees: Morphological and wood traits involved in the balance between weight increase and the gravitropic reaction. J Theor Biol. 2009;256:370–81.
121. Brüchert F, Gardiner B. The effect of wind exposure on the tree aerial architecture and biomechanics of Sitka spruce (Picea sitchensis, Pinaceae). Am J Botany. 2006;93:1512–21.
122. Coutand C, Chevolot M, Lacointe A, Rowe N, Scotti I. Mechanosensing of stem bending and its interspecific variability in five neotropical rainforest species. Ann Botany. 2010;105:341–7.
123. Moulia B, Fournier M. The power and control of gravitropic movements in plants: a biomechanical and systems biology view. J Exp Botany. 2009;60:461–86.
124. Niklas KJ. Plant biomechanics: An engineering approach to plant form and function. Chicago: University of Chicago Press; 1992.

125. Read J, Stokes A. Plant biomechanics in an ecological context. Am J Botany. 2006;93:1546–65.
126. Schopfer P. Biomechanics of plant growth. Am J Botany. 2006;93:1415–25.
127. Schlichting H, Gersten K. Boundary layer theory. 8th ed. Berlin: Springer; 2004.
128. Khir AW, O'Brien A, Gibbs JSR, Parker KH. Determination of wave speed and wave separation in the arteries. J Biomech. 2001;34:1145–55.
129. Mynard J, Penny DJ, Smolich JJ. Wave intensity amplification and attenuation in non-linear flow: Implications for the calculation of local reflection coefficients. J Biomech. 2008;41:3314–21.
130. Mynard JP, Nithiarasu PA. 1D arterial blood flow model incorporating ventricular pressure, aortic valve and regional coronary flow using the locally conservative Galerkin (LCG) method. Comm Num Meth Eng. 2008;24:367–417.
131. Niederer P, Knothe Tate ML, Steck R, Boesiger P. Some remarks on intravascular and extravascular transport and flow dynamics. Int J Cardiovsc Med Sci. 2000;3:21–31.
132. Parker KH, Jones CJH. Forward and backward running waves in the arteries – analysis using the method of characteristics. J Biomech Eng Trans ASME. 1990;112:322–6.
133. Stergiopulos N, Tardy Y, Meister JJ. Nonlinear separation of forward and backward running waves in elastic conduits. J Biomech. 1993;26:201–9.
134. Wang J, Shrive NG, Parker KH, Tyberg JV. "Wave" as defined by wave intensity analysis. Med Biol Eng Comput. 2009;47:189–95.
135. Westerhof BE, van den Wijngaard JP, Murgo JP, Westerhof N. Location of a reflection site is elusive - Consequences for the calculation of aortic pulse wave velocity. Hypertension. 2008;52:478–83.
136. Tang D, Yang C, Kobayashi S, Ku DN. Generalized finite difference method for 3D viscous flow in stenotic tubes with large wall deformation and collapse. Appl Num Math. 2001:38:49–68.
137. Buist M, Sands G, Hunter P, Pullan A. A deformable finite element derived finite difference method for cardiac activation problems. Ann Biomed Eng. 2003;31:577–88.
138. Bathe KJ. Finite element procedures. Upper Saddle River, NJ: Prentice-Hall; 1996.
139. Zienkewicz OC, Taylor RL. Finite element method. 5th ed. Amsterdam: Elsevier; 2000.
140. Huiskes R, Chao EYS. A survey of finite element analysis in orthopedic biomechanics. The first decade. J Biomech. 1983;16:385–409.
141. Crawford NR, Arnett JD, Butters JA, Ferrara LA, Kulkarni N, Goel VK, Duggal N. Biomechanics of a posture-controlling cervical artificial disc: mechanical, in vitro, and finite-element analysis. Neurosurg Focus. 2010;28(6):E11.
142. Doyle BJ, Cloonan AJ, Walsh MT, Vorp DA, McGloughlin TM. Identification of rupture locations in patient-specific abdominal aortic aneurysms using experimental and computational techniques. J Biomech. 2010;43:1408–16.
143. Eriksson A, Nordmark A. Temporal finite element formulation of optimal control in mechanisms. Comp Meth Appl Mech Eng. 2010;199:1783–92.
144. Gantoi FM, Brown MA, Shabana A. AANCF finite element/multibody system formulation of the ligament/bone Insertion site constraints. J Comput Nonlin Dynamics. 2010;5:031006.
145. Hutter R, Schmitt K-U, Niederer P. Mechanical modeling of soft biological tissues for application in virtual reality based laparoscopy simulators. Technol Health Care. 2000;8:15–24
146. Lennon AB, Prendergast PJ, editors. Finite element modelling in biomechanics and mechanobiology. Dublin: Trinity Center for Bioengineering; 2007.
147. Merdji A, Bouiadjra BB, Achour T, Serier B, Chikh BO, Feng ZO. Stress analysis in dental prosthesis. Comput Mat Sci. 2010;49:126–33.
148. Rayfield EJ. Finite element analysis and understanding the biomechanics and evolution of living and fossil organisms. Ann Rev Earth Planet Sci. 2007;35:541–76.
149. Roth S, Raul JS, Willinger R. Finite element modelling of paediatric head impact: Global validation against experimental data. Comput Meth Progr Biomed. 2010;99:25–33.
150. Van Rietbergen B, Weinans H, Huiskes R, Odgaard A. A new method to determine trabecular bone elastic properties and loading using micromechanical finite element models. J Biomech. 1995;28:69–81.
151. Weiss St, Bajka M, Nava A, Mazza E, Niederer P. A finite element model for the simulation of hydrometra. Technol Health Care. 2004;12:259–67.
152. Cohen RCZ, Cleary PW, Mason B. Simulations of human swimming using smoothed particle hydrodynamics. In: Proc. 7th Int. Conf. on CFD in the Minerals and Process Industries. CSIRO, Melbourne, 2009
153. Hieber SE, Koumoutsakos PA. Lagrangian particle method for the simulation of linear and nonlinear elastic model of soft tissue. J Comp Phys. 2008;227:9195–215.
154. Hieber SE, Walther JH, Koumoutsakos P. Remeshed smoothed particle hydrodynamics simulation of the mechanical behavior of human organs. Technol Health Care. 2004;12:305–14.
155. Hoover WG. Smooth particle applied mechanics: the state of the art. River Edge, NJ: World Scientific Publishing; 2006.
156. Koumoutsakos P. Multiscale flow simulations using particles. Ann Rev Fluid Mech. 2005;37:457–87.
157. Yamaguchi T, Ishikawa T, Imai Y, Matsuki N, Xenos M, Deng Y, Bluestein D. Particle-based methods for multiscale modeling of blood flow in the circulation and in devices: challenges and future directions. Ann Biomed Eng. 2010;38:1225–35.
158. Joos G, Freeman IM. Theoretical physics. Mineola, NY: Dover Publications; 1986.
159. Einstein A. Investigations on the theory of Brownian movement. Mineola, NY: Dover Publications; 1956.

160. Piekarski K, Munro M. Transport mechanism operating between blood supply and osteocytes in long bones. Nature.1977;269:80–2.
161. Knothe Tate ML, Niederer P, Knothe U. In vivo tracer transport through the lacunocanalicular system of rat bone in an environment devoid of mechanical loading. Bone. 1998;22:107–17.
162. Knothe Tate ML, Knothe U, Niederer P. Experimental elucidation of mechanical load-induced fluid flow and its potential role in bone metabolism and functional adaptation. Am J Med Sci. 1998;316:189–95.
163. Knothe Tate ML. Whither flows the fluid in bone ? An osteocyte's perspective. J Biomech. 2003;36:1409–24.
164. Swartz MA, Berk DA, Jain RK. Transport in lymphatic capillaries. I. Macroscopic measurements using residence time distribution theory. Am J Physiol Heart Circ Physiol. 1996;270:H324–9.
165. Gashev AA. Physiologic aspects of lymphatic contractile function: current perspectives. Ann New York Acad Sci. 2002;979:178–87.
166. Waterhouse J, Sawdon M, Kirkman E. Capillary dynamics and the interstitial fluid-lymphatic system. Anaesth Int Care Med. 2007;8:73–8.
167. Atkins PW, de Paula J. Physical chemistry. Oxford: Oxford University Press; 2001.
168. Landau LD, Lifshitz EM. A course of theoretical physics. Vol. 5. Ed 4. Oxford: Butterworth-Heinemann; 1980
169. Michel CC, Phillips ME. Fluid exchange in the microcirculation. J Physiol. 1987;388:421–35.
170. Hu X, Weinbaum S. A new view of Starling's hypothesis at the microstructural level. Microvasc Res. 1999;58:281–304.
171. Van den Berg B, Vink H, Spaan JA. The endothelial glycocalyx protects against myocardial edema. Circ Res. 2003;92:592–4.
172. Renkin EM. Some consequences of capillary permeability to macromolecules: Starling's hypothesis reconsidered. Am J Physiol Heart Circ Physiol. 1986;250:H706–10.
173. Curry FRE. Microvascular solute and water transport. Microcirculation. 2005;12:17–31.

TABLE OF CONTENTS

Critical Reviews™ in Biomedical Engineering

Volume 38

SUBJECT INDEX – Volume 38

Critical Reviews™ in Biomedical Engineering

Page Numbers for Issues:

Issue 1, 1–116; **Issue 2,** 117–224; **Issue 3,** 225–304; **Issue 4,** 305–392; **Issue 5,** 393–486; **Issue 6,** 487–578

AUTHOR INDEX – Volume 38

Critical Reviews™ in Biomedical Engineering

Page Numbers for Issues:

Issue 1, 1–116; **Issue 2,** 117–224; **Issue 3,** 225–304; **Issue 4,** 305–392; **Issue 5,** 393–486; **Issue 6,** 487–578