

Physical methods for structure, dynamics and binding in immunological research

Dimitrios Morikis¹ and John D. Lambris^{2,*}

¹Department of Chemical and Environmental Engineering, University of California at Riverside, Riverside, CA 92521, USA

²Protein Chemistry Laboratory, Department of Pathology and Laboratory Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA

We present four experimental physical methods – X-ray and neutron diffraction, nuclear magnetic resonance spectroscopy, mass spectrometry and calorimetry – and two computational methods – molecular dynamics simulations and electrostatics calculations – which are general and widely applicable in the study of protein structure, dynamics and binding. These methods are useful tools for biologists that lead to structure–function, dynamics–function and binding–function correlations, in efforts to understand biomolecular function. Standard and emerging technologies within these methods are discussed and representative examples of applications in immunology are presented, from antigen–antibody, complement and MHC–T-cell receptor research. The examples demonstrate the power of the reviewed methods in immunological studies at the molecular level.

The post-genomic era is characterized by an abundance of linear sequences and by an increasing number, although still limited, of three-dimensional protein structures. A typical progression of discovery follows the order of (i) sequence determination of whole genomes for various species, using high-throughput methods, (ii) proteomics analyses using high-throughput technologies to identify protein subcellular localization and protein–protein interactions, and for functional studies, (iii) structure determination of single proteins and protein–protein, protein–nucleic acid, protein–ligand, protein–carbohydrate, enzyme–substrate, protein–drug, protein–metal, protein–solvent, protein–lipid, ternary and higher order complexes, (iv) study of protein dynamics, (v) study of binding, including kinetics and thermodynamics, and (vi) elucidation of structure–dynamics–binding–function correlations. Hypothesis-driven discovery based on known facts is also part of this database-driven research. Despite intense research in the area of genomics and bioinformatics, knowledge of sequence does not so far guarantee accurate prediction of function.

Corresponding authors: Dimitrios Morikis (dmorikis@engr.ucr.edu), John D. Lambris (lambris@mail.med.upenn.edu).

* John D. Lambris has been a member of the *Trends in Immunology* Editorial Board since 2001.

Available online 8 October 2004

Proteins interact with other proteins, nucleic acids, ligands, prosthetic groups, carbohydrates, lipids, metal ions and inorganic molecules, and when the proteins are enzymes, with substrates and co-factors (we invariably call them protein-complexes hereafter). Proteins and protein complexes are dynamic entities participating in a wide range of motions and their dynamics are essential for function. The interaction of proteins and protein complexes with solvent molecules mediates several of their dynamic motions.

With the availability of complete genomes from various species, the need for structural genomics is obvious. Some attempts have been made at high-throughput methods [1–5]. The purpose of this Review is to provide an anthology of ‘low-throughput’ physical methodologies that are useful for studying protein structure, dynamics and interactions, drawing examples from immunology. The experimental and computational physical methods we review here are either established or show high promise as aids in the quest to understand immune system function at a molecular level. We define immunophysics, in general, as the underlying physics responsible for structure, dynamics and interactions, which describe immune system function at atomic, molecular and cellular level. The examples have caught our attention they are only representative, from a large literature pool, addressing three major areas of immunology, antigen–antibody, complement and MHC–T-cell receptor (TCR) research.

Protein crystallography

X-ray diffraction is the classic method for atomic resolution structure determination of proteins and protein complexes in a single crystal state (Box 1). At the end of June 2004, 21 531 out of 24 887 depositions in the protein data bank (PDB) [6] are protein structures, including complexes, determined by X-ray diffraction. The remaining are protein structures determined by nuclear magnetic resonance (NMR) spectroscopy, neutron diffraction and low-resolution techniques, such as electron microscopy, electron diffraction and fiber diffraction. A brief discussion on the capabilities, practices, limitations, dilemmas and emerging technologies for crystallographic

Box 1. Protein crystallography

Physical effect: diffraction of X-rays or neutrons by biomolecules arranged in a lattice.

X-ray diffraction

Direct measurements: diffraction patterns.

Derived parameters: electron scattering densities.

Sample state: single crystal phase.

Sample condition or perturbation: low temperature.

Applications:

- Structure determination of proteins and protein complexes
- Thermal fluctuations, flexibility, and mobility

Neutron diffraction

Direct measurements: diffraction patterns.

Derived parameters: nuclear scattering densities.

Sample state: single crystal phase; D₂O solutions, variable H₂O/D₂O mixtures, deuterated samples for H/D diffraction differences.

Sample conditions or perturbations: H/D exchange; low temperature.

Applications:

- Structure determination of proteins and protein complexes
- Dynamics

methods has been recently published (see editorial overview and references therein [7]).

Given that the alternative method for atomic resolution structure determination, NMR spectroscopy (Box 2), has a molecular size limitation and is slower than X-ray diffraction (Box 3), it is not surprising that increasing numbers of large and multi-component complexes appearing in the PDB are solved by X-ray diffraction. This is in part owed to the use of synchrotron radiation facilities, which generate high energy tunable X-rays, and the use of the multiwavelength anomalous diffraction (MAD) method, typically in combination with selenium labeling. Synchrotron radiation produces sharper diffraction patterns and allows for significant reduction of the data collection time. Selenium-labeled samples can be generated in an efficient manner by protein expression prior to crystallization. The use of selenium eliminates the need for several crystals doped with different heavy metals, previously required to solve the phases of the diffraction patterns [8].

Determination of structure using X-ray diffraction also provides an assessment of the dynamic character, although not a direct measurement, of the protein or the protein complex. The temperature factor measures the atomic conformational and crystal lattice disorder and thermal vibration. Conformational disorder of flexible regions, such as loops or extended coils, or structured regions, which are mobile with respect to the rest of the protein, result in reduced and sometimes lost electron density. However, NMR studies provide direct quantitative measures of biomolecular dynamics and internal mobility.

Another source for crystallography studies are neutrons as opposed to X-rays (Boxes 1 and 3). Neutron diffraction studies can provide the atomic coordinates of hydrogen atoms, which are not provided by X-ray diffraction at regular resolutions [9,10]. At the PDB [6] there are several impressive examples of crystallographic structures of large multi-component complexes, including complexes with

Box 2. Nuclear magnetic resonance (NMR) spectroscopy

Physical effect: measurement of magnetic resonance of spin-possessing nuclei in a magnetic field with applied radiofrequency waves; magnetization transfer among interacting nuclei.

Liquid state NMR spectroscopy (conventional, multidimensional and multinuclear)

Direct measurement: chemical shifts, linewidths, spin-spin J-coupling constants, nuclear Overhauser effects (NOEs), peak or cross peak intensities, areas or volumes.

Derived parameters: distances, torsion angles, chemical shift indices, dissociation constants, T₁ and T₂ relaxation times, order parameters, H/D exchange rates, protection factors.

Deduced parameters: hydrogen bonds.

Sample state: solution phase; sophisticated experiments require ¹³C- and/or ¹⁵N-labeled sample to increase spectral resolution.

Sample conditions or perturbations: uniform or selective ¹³C- and/or ¹⁵N-isotopic labeling; deuteration; H/D exchange; variation in temperature or pH; titration of one sample component in case of complexes.

Applications:

- Secondary structure determination
- Structure determination for proteins
- Structure determination for protein complexes
- Protein-protein, protein-ligand interaction and protein multimerization, without complete structure determination
- Dynamics of proteins and protein complexes at atomic resolution
- Protein folding
- Specialized uses of liquid state NMR for the study of side chain dynamics, hydration, slow exchange aromatic ring flips, proline *cis-trans* isomerization, histidine tautomerization, pK_a values of ionizable residues, paramagnetic effects.

TROSY (transverse relaxation-optimized spectroscopy) NMR

Direct measurement: ¹H-¹⁵N and aromatic ¹H-¹³C resonances.

Sample state: solution phase; perdeuterated or at least deuterated sample to ~70% and ¹⁵N- and/or ¹³C-isotopically labeled sample; tailored perdeuteration may be desired for certain applications.

Sample conditions or perturbations: same as in standard liquid state NMR.

Applications:

- Aid in resonance assignments for proteins of molecular mass up to or > 100 kD
- Identification of interface in protein complexes using chemical shift perturbation mapping or hydrogen exchange
- Detection of hydrogen bonds

Liquid crystal state NMR spectroscopy

Direct measurement: residual dipolar couplings (RDCs).

Derived parameters: orientational restraints.

Sample state, conditions, or perturbations: solution state in presence of liquid crystal media, dilute aqueous phospholipid liquid crystal solutions forming bicelles (disk-like lipid bilayers), filamentous bacteriophages or polyacrylamide gels; ¹⁵N- or ¹³C-isotopically labeled proteins; temperature variation.

Applications:

- Study of multidomain proteins, protein-membrane complexes and protein complexes

glycans and complexes with lipids, which imitate the membrane environment.

Examples (Table 1) of use in antigen-antibody research are the structures of several human anti-HIV-1 gp120-reactive antibodies and ternary complexes of fragments of human HIV-1 gp120 envelope glycoproteins complexed with fragments of CD4 and fragments of an induced neutralizing antibody [11,12]. In complement research, a systematic effort to piece together the components of C1

Box 3. Method comparisons**Neutron vs X-ray crystallography**

Neutron crystallography: provides hydrogen atom positions; diffraction maps are sharper but neutron facilities are scarce; requires long exposure times and large crystals.

X-ray and neutron diffraction vs nuclear magnetic resonance (NMR) spectroscopy

X-ray and neutron diffraction limitations: (i) Not all proteins or protein complexes are amenable to crystallization. (ii) Crystalline state is non-physiological. (iii) Crystalline state is quasi-static, which means with significant loss of dynamic character compared to solution. This might not always be a limitation because cell or membrane interior might also restrict protein dynamics. (iv) Data are often collected at low temperature, which is a non-physiological condition. (v) Positions of hydrogen atoms are not provided at regular resolutions (X-ray only).

X-ray and neutron diffraction advantages: (i) No molecular mass limit. (ii) Faster for data collection with more automated data analysis.

NMR limitations: (i) Complete structure determination is possible for molecular mass up to ~30 kD but most structures determined by NMR fall in the molecular mass range 3–14 kD with the distribution peaking at ~8 kD. However, the TROSY (transverse relaxation-optimized spectroscopy) method enables spectral collection of proteins with molecular mass more than >30 kD for up to ~1 MD. (ii) Isotopic labeling is usually necessary for conventional NMR. (iii) Perdeuteration and isotopic labeling is necessary for TROSY NMR. (iv) Analysis is time-consuming despite automation efforts.

NMR advantages: (i) Provides positions for hydrogen atoms. (ii) Enables studies of dynamics, folding and other specialized measurements. (iii) Study of membrane-bound proteins with liquid crystal or solid state NMR.

ESI vs MALDI MS

Molecular mass limits: ESI MS: up to 100 kD; MALDI MS up to 1 MD.

Sample concentration: ESI MS, femtomole-picomole (nanoESI MS: zeptomole-femtomole); MALDI MS, femtomole.

ESIMS advantage: ESI MS is optimal for HPLC/MS on-line coupling.

ITC vs DSC calorimetry

ITC and DSC are complementary to provide a complete set of thermodynamic parameters.

(recognition molecule C1q and serine proteases C1r and C1s) has been presented together with the role of Ca²⁺ ions in holding together the various components of the complex [13,14]. Based on these studies, models of the heterotrimeric globular domain assembly of C1q in complex with IgG1 and with the C-reactive protein have been constructed. In MHC–TCR research, the structures of a human HLA class I MHC protein fragment in complex

with a microglobulin fragment, a peptide and a TCR fragment [15], and the structure of a human cytomegalovirus protein fragment in complex with an HLA class I MHC protein fragment, a microglobulin fragment and a T-cell lymphotropic virus peptide [16], have been reported.

Efforts for high-throughput crystallography have been implemented including new and automated crystal growth procedures and robotics for data collection [1–4]. More progress is expected in this area; however, the current rate of structure depositions at the PDB [6] suggests that high-throughput crystallography is still at the beginning. This is mainly because of the complexity of the crystallographic methods, which involve several steps from sample preparation, to data collection, to analysis, followed by computational fitting and modeling.

NMR spectroscopy

Liquid state NMR spectroscopy [17] has been an efficient method for the structure determination of small proteins (<30 kD) in solution at atomic resolution [18] (Boxes 2 and 3). For peptides and very small proteins, two-dimensional ¹H–¹H homonuclear NMR studies are sufficient for structure determination. For larger proteins, two-dimensional and three-dimensional (and occasionally four-dimensional) NMR spectra are measured using ¹⁵N- and/or ¹³C-isotopically labeled samples, involving interactions between or among ¹H, ¹⁵N and ¹³C nuclei. In two-dimensional NMR, each dimension corresponds to chemical shifts of different resonances, (e.g. ¹H, ¹⁵N, ¹³C) and pairwise interaction between two nuclei is represented by cross peaks, (e.g. ¹H–¹H, ¹H–¹⁵N, ¹³C–¹⁵N) in two-dimensional maps. In three-dimensional NMR, three-way interactions (e.g. ¹H–¹⁵N–¹³C) are represented by cross peaks in three-dimensional maps [19,20]. Increase in spectral dimensionality results in increased spectral resolution and simplification of the spectral assignments. Interactions can be through-bond, involving scalar couplings, or through-space, involving dipolar couplings. In particular, heteronuclear triple resonance ¹H–¹⁵N–¹³C NMR spectra in combination with automated assignments procedures [21] have resulted in increased efficiency for structure determination. Significant improvements in spectral sensitivity have been made possible with recent advances in superconducting magnet technology (<900 MHz for protons) and with cryogenic probe technologies (cryoprobes) [19].

Three-dimensional structure determination is possible using restrained computational methods (e.g. distance geometry and molecular dynamics-based simulated annealing) with NMR-derived or NMR-deduced restraints [22]; these are nuclear Overhauser effect (NOE)-derived distance restraints, J-coupling constant-derived torsion angle restraints (sometimes aided by NOEs), direct J-coupling constant restraints, chemical shift restraints and hydrogen/deuterium (H/D) exchange-deduced distance-type backbone hydrogen bond restraints (aided by NOEs and J-coupling constants). Without computational methods, secondary structure determination is possible by analysis of chemical shift indices (CSIs), NOE-connectivity patterns, ³J_{HN–H α} -coupling constants and slow backbone amide H/D exchange.

Table 1. References for representative immunological examples

Method	Antigen-antibody research	Complement research	MHC–T-cell receptor research
Crystallography	[11,12]	[13,14]	[15,16]
Nuclear magnetic resonance	[32,33]	[34,35]	[36,37]
Mass spectrometry	[49,50]	[51,52]	[53,54]
Calorimetry	[64,65]	[66], (M. Katragadda <i>et al.</i> , unpublished)	[67,68]
Computational modeling	[78,79]	[80,81]	[82,83]

Structure determination by NMR is possible for tightly bound protein complexes (in slow exchange) using isotopic labeling and special isotopically filtered experiments [23], and for weakly bound (in fast exchange) small ligands to large proteins (but not complexes) using spectra of exchange NOEs called transfer NOEs (trNOEs) [24]. In addition, if spectral assignments of each component of the complex are available, chemical shift perturbation mapping studies can identify the binding site by observing chemical shift differences when comparing spectra of the free molecules and the complex [19]. Dissociation constants can also be determined in chemical shift titration experiments of variable ligand–target ratios.

NMR has been widely used to study protein backbone and side-chain dynamics on a wide range of timescales [25,26]. One of the most popular studies is the backbone dynamics at picosecond–nanosecond time scale, using amide groups. These provide a measure of the diffusional motion (or fluctuation) of the ^1H – ^{15}N bond, the time scale of the motion and the flexibility of the protein as whole. Typical NMR parameters are T_1 and T_2 relaxation times, ^1H – ^{15}N heteronuclear NOE and order parameters.

NMR in combination with H/D exchange has been used to study the internal mobility (segmental or global at millisecond to megasecond time scale), packing, local stability and folding of proteins [27]. Slow H/D exchange rates occur for amides that are protected from solvent, either because they are in the interior of proteins or because they are in the middle of well-defined elements of regular hydrogen-bonded secondary structure. A useful parameter is the calculation of protection factors. For protein folding, H/D exchange is typically used in combination with rapid mixing and pulse-labeling methods. Folding studies use secondary structure analysis and the determination of tertiary contacts.

Structural studies by NMR, but not complete structure determination, have been possible for proteins with a molecular mass up to or greater than 100 kD, using the TROSY (transverse relaxation-optimized spectroscopy) experiment and perdeuterated samples [28]. The TROSY effect has been incorporated in heteronuclear triple resonance NMR experiments.

Standard or special NMR experiments and analyses can be used for specialized studies summarized in Box 2.

Liquid crystal state NMR of proteins and protein complexes enables the measurement of residual dipolar couplings (RDCs), which correspond to orientational restraints [29]. Liquid crystal state NMR is based on the property of certain media to align when placed into magnetic fields. Aligning media, such as dilute bicelle solutions, filamentous bacteriophages or polyacrylamide gels, form physical barriers for the rotational tumbling of proteins, which is restricted to being anisotropic, thus enabling the measurement of RDCs.

The structure determination of membrane proteins has been problematic because they are difficult to crystallize for study by X-ray or neutron diffraction and they aggregate in solution for study by NMR. This is a result of the large hydrophobic character of their transmembrane regions, which become exposed in the absence of membrane. Liquid crystal state NMR methods are also efficient

for the study of small membrane proteins and peptides bound in bicelles or micelles because bicelles, micelles and membranes share the characteristic that they are formed by lipid bilayers [30]. Solid-state NMR methods have also been used to study proteins bound to bicelles [30] or amyloid fibrils [31], however, these methods will not be reviewed here.

Examples (Table 1) of NMR-derived structures of a mouse anti-HIV antibody complexed with the gp120 V3 peptide [32] and a V3 peptide bound to human anti-HIV-1 neutralizing antibody [33] have been reported. In complement research, examples of NMR studies of multi-domain proteins are the structure and dynamics of human regulators of complement activation, a component of decay accelerating factor (DAF; CD55) [34] and fragments of complement receptor 1 (CR1; CD35) [35]. In MHC–TCR research, the structure and dynamics of a fragment of mouse single chain TCR [36] and the structure of a domain of human MHC protein [37] have been reported.

Structural genomic efforts for high-throughput NMR have been implemented [4,5], however, the complexity of the NMR studies so far did not allow for much progress. NMR lags behind X-ray studies in automation because of the multiplicity and diversity of required experiments, the complexity of data analysis and the inherent molecular size limitation.

Mass spectrometry (MS)

MS (Box 4) has been used to determine protein and peptide molecular mass and to provide sequence information [38,39]. Sequence determination by MS, also known as *de novo* protein sequencing, is useful in the characterization of post-translational protein modifications (e.g. phosphorylation, methylation, glycosylation, myristylation). In this sense, MS is complementary to sequence deduction of unknown proteins from genomics studies, which are based on translating nucleotide sequence to amino acid sequence. Major technological advances for ‘soft’ ionization in electrospray ionization (ESI) MS [40] and matrix-assisted laser desorption ionization (MALDI) MS [41], combined with ion trap (IT), time-of-flight (TOF), quadrupole (Q), triple quadrupole or quadrupole-TOF (Q-TOF) mass analyzers [42], have made the study of several biological processes possible.

Several protocols to obtain sequence information of an unknown protein have been reported [38,43,44]. Typically, protein MS is combined with protein cleavage into peptide fragments. Cleavage by different specific proteases is usually necessary and the peptides are separated by reversed-phase high performance liquid chromatography (HPLC). Groups of peptides are fed on-line to the MS spectrometer. Sequence information for the peptides can be obtained using tandem MS (MS/MS), in which initial ionization is followed by secondary ionization to smaller ions, for example, by collision-induced dissociation (CID) [42]. The final step is to piece together the sequences of the cleavage peptides to form the protein sequence, with the aid of database searches. ESI MS, or combination of MALDI and ESI MS, with variations in the choice for secondary ionization and mass analysis method, can be used [38,39].

Box 4. Mass spectrometry (MS)

Physical effect: sample ionization and measurement of mass/charge ratio.

Electrospray ionization (ESI) MS; matrix-assisted laser desorption ionization (MALDI) MS

Direct measurement: mass:charge ratio of ionized peptide, proteins and fragmented proteins into peptides.

Derived parameters: mass; protein sequence information.

Sample state: solution phase; tissues.

Sample conditions or perturbations: fragmentation of proteins into peptides; H/D exchange; variations in pH or ionic strength; isotopic labeling; cross-linking.

Applications:

- Determination of peptide and protein molecular mass
- Proteomics analysis for protein identification and characterization of post-translational modifications
- Identification of components of protein complexes, multimerization state and stoichiometry; aggregation and fibril formation
- Identification of interface of protein complexes
- Real time dynamics of protein–protein interactions or subunit exchange in multimeric proteins
- Kinetic profiles of binding
- Protein folding
- Tissue imaging

MS has been used to identify the components of protein complexes, the oligomerization state of proteins, the stoichiometry of multi-subunit proteins or complexes and their changes as a function of pH, temperature, salt and ligand binding [39]. MS in combination with H/D exchange has been used to study protein complexes, protein dynamics, kinetics and protein folding [45,46]. The principles for hydrogen–deuterium exchange studies by MS are the same as for NMR, with the exception that deuterons are silent in proton NMR. Also, MS provides segment-specific information whereas NMR provides site-specific information. NMR can provide more complete structural insights if the three-dimensional structure is available because of its site specificity. However, efficient NMR spectroscopy requires isotopic labeling, whereas the process of complete resonance assignments of the NMR spectra is slow. MS has the advantage that data collection is faster and enables studies of a larger molecular mass range, compared to NMR.

New breakthrough imaging technologies of MS for the study of the topology of proteins in tissues have emerged (reviewed in Ref. [47]). Also, combination of surface plasmon resonance (SPR) with MS (SPR–MS) has been reported for the study of protein complexes [48].

ESI or MALDI MS have been used to study molecular interactions in antigen–antibody complexes [49,50]. Examples from complement research are the identification of the interaction site of C3b with Factor B [51] and the demonstration of post-translational C-mannosylation in the four terminal components of the complement system, C6–C9 [52]. In MHC–TCR research, ESI or MALDI MS methods have been developed to study mixtures of MHC protein–peptide complexes [53,54]. An educational review on mass spectrometry and structural immunology has been reported [55].

Various studies including proteomic analyses using MS, some including ‘top-down’ and ‘bottom-up’ approaches, have been reviewed elsewhere (e.g. Refs [38,56,57]).

Calorimetry

Isothermal titration calorimetry (ITC) (Box 5) is a method to study the thermodynamics of protein complex formation [58–62]. ITC performs a precise measurement of the heat released by a biochemical reaction as a function of time. Heat is then converted into binding enthalpy (ΔH) and binding constant K_b . The latter is used to calculate the binding free energy (ΔG). By performing the ITC experiments at different temperatures, heat capacity can be calculated from enthalpy. The advantage of ITC over other methods that are used to measure or calculate K_b s or ΔG s, is that ITC distinguishes the components of ΔG , enthalpy and entropy (ΔS , where $\Delta G = \Delta H - T\Delta S$). This is important because there are several combinations of ΔH and ΔS that can produce the same ΔG , and enthalpy and entropy changes are due to different chemical and structural processes. For example, enthalpy changes occur mainly because of physico–chemical interactions between the two components of the complex, such as van der Waals, electrostatic and hydrogen bonding, compared to interactions with the solvent. Entropy changes occur because of the removal of solvent molecules from the binding interface (desolvation), which results in increased solvent entropy, and because of a loss of conformational freedom for residues located at the binding interface. These processes are accompanied by structural rearrangements. Although ITC is not suitable thus far to measure strong binding (K_b of $\sim 10^9 \text{ M}^{-1}$ or higher), it can still establish a lower limit and can quantitate enthalpy changes. In some

Box 5. Calorimetry

Physical effect: measurement of reaction heat as a function of time or temperature.

Isothermal titration calorimetry (ITC)

Direct measurement: electrical power (heat/time) that is required to equilibrate the temperature on binding in a reaction cell with respect to a reference cell, as a function of time.

Derived parameters: heat difference (ΔQ), enthalpy (ΔH), binding constant (K_b), stoichiometry (n), binding free energy (ΔG), entropy (ΔS), heat capacity at constant pressure (ΔC_p).

Sample state: solution phase; titration of one component of the complex; concentration variation.

Sample conditions or perturbations: variations in pH, ionic strength, temperature, buffer composition, solvent viscosity.

Applications:

- Thermodynamics of binding
- Enzymatic kinetics

Differential scanning calorimetry (DSC)

Direct measurement: difference in apparent excess heat capacity of the protein in solution and the solution alone (ΔC_p)_{app}, as a function of temperature, in a protein that undergoes a transition.

Derived parameters: ΔC_p , ΔH , transition temperature (T_m).

Sample state: solution phase.

Sample conditions or perturbations: temperature variation; pressure variation.

Applications:

- Protein stability and folding

favorable cases, changes in temperature, pH and solvent composition can alter the binding constant to the range of applicability of ITC [62].

Another microcalorimetry method is differential scanning calorimetry (DSC) (Boxes 3 and 5) and its variation temperature-modulated DSC, which are used for the study of protein stability and protein-folding transitions as a function of temperature [58,61,63].

ITC in combination with X-ray crystallography has been used to study antigen–antibody interactions [64,65]. In complement research, ITC has been used to study the thermodynamics of binding of complement inhibitor peptide compstatin to C3 (M. Katragadda *et al.*, unpublished); DSC in combination with NMR, CD, analytical ultracentrifugation, fluorescence spectroscopy and computational modeling has been used to study the folding of complement inhibitor vaccinia control protein (VCP), a homolog of CR1 [66]. In MHC–TCR research, ITC and SPR have been used to study the thermodynamics and kinetics of TCR–peptide–MHC protein interactions [67] and DSC has been used to study the stability of a MHC protein–antigenic peptide complex [68].

ITC can be creatively used in combination with protein crystallography, NMR, MS, SPR and computational studies for binding studies. High-throughput uses of calorimetry have also been initiated [58].

Computational modeling

Computational modeling using biophysical principles is an integral part of biology and has been used to understand experimental data and to predict biophysical properties related to protein structure, dynamics and interactions ([69–71], see editorial overviews and references therein).

The three-dimensional structures determined using protein crystallography or NMR spectroscopy provide an overall time-averaged or ensemble-averaged topology, which is useful but does not fully represent the dynamic character of proteins and protein complexes. Experimental studies using different types of spectroscopy demonstrate the presence of a wide variety of atomic, local, segmental or global motions in biomolecules, at a large range of time frames (femtoseconds to megaseconds). Often processes, such as protein or ligand binding and enzymatic catalysis, cannot be understood in terms of structures alone. Structural rearrangements and dynamic motions are typically involved in intra- and inter-molecular interactions.

Classical molecular dynamic (MD) computer simulations (Box 6), based on experimentally determined structures, typically provide information for processes in the picosecond to microsecond time scale. Besides understanding the native motions of biomolecules, MD simulations have been used to understand biological processes, such as folding, conformational transitions, protein or ligand binding, enzymatic catalysis, biological reactions and protein–membrane interactions [72,73].

Classical electrostatic calculations (Box 6) are a significant aid to understand biological processes, when experimentally determined structures are available. Electrostatic interactions are ubiquitous in structure and dynamics of single proteins or protein complexes [74,75].

Box 6. Computational studies

Molecular dynamics (MD) simulations

Physical effect: numerical solution of Newton's equation of motion using a molecular mechanics force field.

Calculated quantities: MD trajectory (evolution of conformation of biomolecule or complex as a function of time); potential energy.

Derived quantities: time scales of motions or structural transitions, free energies, motional amplitudes.

Sample state: previously determined protein structure by experimental methods.

Sample conditions or perturbations: variation of thermodynamic ensemble; variation of temperature; conformational constraints; solvent representation; theoretical mutations or chemical changes.

Applications:

- Fluctuations and dynamics
- Conformational transitions, folding and stability
- Interactions of protein domains, modules or components of complexes
- Enzymatic catalysis
- Preparation for other theoretical studies (e.g. docking or electrostatic calculations)
- Structure refinement
- Protein–membrane interactions and solvent penetration into membranes

Electrostatic calculations

Physical effect: solution of Poisson-Boltzmann or Coulomb's equation.

Calculated quantity: electrostatic potential.

Derived quantities: free energies, binding constants, apparent pK_a values.

Sample state: previously determined protein structure by experimental methods.

Sample conditions or perturbations: variations in pH, ionic strength, or temperature; theoretical mutations or chemical changes.

Applications:

- Protein complex formation
- Stability
- Solvation properties
- Ionization properties
- Enzymatic catalysis

Short-range interactions contribute to the structure specificity and affect the stability of the biomolecule, whereas long-range interactions in solution contribute to recognition and binding. Electrostatic interactions are due to the presence of electric dipoles within a biomolecule or because of charges possessed by ionizable amino acids (and other groups present) in their ionized form. Electrostatic interactions are also used to understand solvation effects, to model ionization, to predict apparent pK_a values and to understand protein complex formation and stability. Modeling of electrostatic interactions is a significant component of automated protein–protein or protein–ligand docking algorithms [76]. Creative studies using combination of MD simulations and electrostatic calculations have been reported [77].

Examples (Table 1) of antigen–antibody research using electrostatic calculations are the prediction of pK_a values of ionizable groups of an antigen–antibody complex [78] and the docking of antigens to antibodies [79]. In complement research, MD simulations of complement inhibitor peptide compstatin revealed an ensemble of inter-converting conformers in solution [80] and electrostatic calculations proposed that the C3d–complement receptor 2 (CR2) recognition is driven by long-range

electrostatic interactions [81]. Electrostatic interactions have been used to study various MHC protein–peptide interactions [82] and MD simulations have been used to determine free energies of binding and the recognition mechanism in the formation of two TCR–immunogenic peptides–MHC protein complexes [83].

Perspectives

The physical methods we have presented aim to elucidate the determination of biomolecular structure, dynamics and binding. Biomolecules communicate to form complexes and dynamics are important for function because they aid communication and enable access to binding sites. All experimental and computational physical methods presented here measure or provide insights into complex formation. The technical aspects of these methods have improved to enable the study of larger systems in a faster way. Crystallographic studies have produced spectacular examples of large multi-component complexes. The same is true for MS. NMR spectroscopy is reaching for larger molecular mass systems and at a wider range of dynamics, in addition to several specialized studies. ITC is promising to become the standard method to measure the thermodynamics of binding. The computational methods, molecular dynamics simulations and electrostatic calculations, have benefited from advances in computer processing power and memory capacities and from novel physical formulations, which enable more realistic simulations. Innovations show promise for studies of interactions of proteins with membranes or membrane components. Finally, several efforts have been initiated for high-throughput studies using most of the physical methods presented here.

Acknowledgements

Support by NIH (grants GM-62134, AI-30040) and the American Heart Association, Western States Affiliate (grant-in-aid 0255757Y) is gratefully acknowledged.

References

- Heinemann, U. *et al.* (2003) Facilities and methods for the high-throughput crystal structural analysis of human proteins. *Acc. Chem. Res.* 36, 157–163
- Adams, M.W.W. *et al.* (2003) The Southeast Collaboratory for Structural Genomics: A high-throughput gene to structure factory. *Acc. Chem. Res.* 36, 191–198
- Kyogoku, Y. *et al.* (2003) Structural genomics of membrane proteins. *Acc. Chem. Res.* 36, 199–206
- Yee, A. *et al.* (2003) Structural proteomics: toward high-throughput structural biology as a tool in functional genomics. *Acc. Chem. Res.* 36, 183–189
- Staunton, D. *et al.* (2003) NMR and structural genomics. *Acc. Chem. Res.* 36, 207–214
- Berman, H.M. *et al.* (2002) The Protein Data Bank. *Acta Crystallogr. D Biol. Crystallogr.* 58, 899–907
- Moffat, K. and Chait, B.T. (2003) Biophysical methods: doing more with less. *Curr. Opin. Struct. Biol.* 13, 535–537
- Hendrickson, W.A. (2000) Synchrotron crystallography. *Trends Biochem. Sci.* 25, 637–643
- Shu, F. *et al.* (2000) Enhanced visibility of hydrogen atoms by neutron crystallography on fully deuterated myoglobin. *Proc. Natl. Acad. Sci. U. S. A.* 97, 3872–3877
- Gabel, F. *et al.* (2002) Protein dynamics studied by neutron scattering. *Q. Rev. Biophys.* 35, 327–367
- Huang, C.C. *et al.* (2004) Structural basis of tyrosine sulfation and V_H-gene usage in antibodies that recognize the HIV type 1 coreceptor-binding site on gp120. *Proc. Natl. Acad. Sci. U. S. A.* 101, 2706–2711
- Kwong, P.D. *et al.* (1998) Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody. *Nature* 393, 648–659
- Gaboriaud, C. *et al.* (2004) Structure and activation of the C1 complex of complement: unraveling the puzzle. *Trends Immunol.* 25, 368–373
- Gaboriaud, C. *et al.* (2003) The crystal structure of the globular head of complement protein C1q provides a basis for its versatile recognition properties. *J. Biol. Chem.* 278, 46974–46982
- Stewart-Jones, G.B.E. *et al.* (2003) A structural basis for immunodominant human T cell receptor recognition. *Nat. Immunol.* 4, 657–663
- Gewurz, B.E. *et al.* (2001) Antigen presentation subverted: structure of the human cytomegalovirus protein US2 bound to the class I molecule HLA-A2. *Proc. Natl. Acad. Sci. U. S. A.* 98, 6794–6799
- Ernst, R.R. (1992) Nuclear magnetic resonance Fourier transform spectroscopy (Nobel lecture). *Angew. Chem.-Int. Ed. Engl.* 31, 805–930
- Wuthrich, K. (2003) NMR studies of structure and function of biological macromolecules (Nobel Lecture). *Angew. Chem.-Int. Ed Engl.* 42, 3340–3363
- Ferentz, A.E. and Wagner, G. (2000) NMR spectroscopy: a multifaceted approach to macromolecular structure. *Q. Rev. Biophys.* 33, 29–65
- Clore, G.M. and Gronenborn, A.M. (1998) Determining the structures of large proteins and protein complexes by NMR. *Trends Biotechnol.* 16, 22–34
- Moseley, H.N.B. *et al.* (2001) Automatic determination of protein backbone resonance assignments from triple resonance nuclear magnetic resonance data. In *Nuclear Magnetic Resonance of Biological Macromolecules, Methods in Enzymology Pt B* 339, 91–108
- Guntert, P. (1998) Structure calculation of biological macromolecules from NMR data. *Q. Rev. Biophys.* 31, 145–237
- Breeze, A.L. (2000) Isotope-filtered NMR methods for the study of biomolecular structure and interactions. *Prog. Nucl. Magn. Reson. Spectrosc.* 36, 323–372
- Post, C.B. (2003) Exchange-transferred NOE spectroscopy and bound ligand structure determination. *Curr. Opin. Struct. Biol.* 13, 581–588
- Ishima, R. and Torchia, D.A. (2000) Protein dynamics from NMR. *Nat. Struct. Biol.* 7, 740–743
- Kay, L.E. (1998) Protein dynamics from NMR. *Nat. Struct. Biol.* 5, 513–517
- Dempsey, C.E. (2001) Hydrogen exchange in peptides and proteins using NMR-spectroscopy. *Prog. Nucl. Magn. Reson. Spectrosc.* 39, 135–170
- Fernandez, C. and Wider, G. (2003) TROSY in NMR studies of the structure and function of large biological macromolecules. *Curr. Opin. Struct. Biol.* 13, 570–580
- Prestegard, J.H. (1998) New techniques in structural NMR – anisotropic interactions. *Nat. Struct. Biol.* 5, 517–522
- Opella, S.J. *et al.* (2001) Nuclear magnetic resonance of membrane-associated peptides and proteins. In *Nuclear Magnetic Resonance of Biological Macromolecules, Methods in Enzymology Pt B (Vol. 339)*, pp. 285–313
- Tycko, R. (2001) Solid-state nuclear magnetic resonance techniques for structural studies of amyloid fibrils. In *Nuclear Magnetic Resonance of Biological Macromolecules, Methods in Enzymology Pt B (Vol. 339)*, pp. 390–413
- Tugarinov, V. *et al.* (2000) NMR structure of an anti-gp120 antibody complex with a V3 peptide reveals a surface important for co-receptor binding. *Structure* 8, 385–395
- Sharon, M. *et al.* (2003) Alternative conformations of HIV-1V3 loops mimic β hairpins in chemokines, suggesting a mechanism for coreceptor selectivity. *Structure* 11, 225–236
- Uhrinova, S. *et al.* (2003) Solution structure of a functionally active fragment of decay-accelerating factor. *Proc. Natl. Acad. Sci. U. S. A.* 100, 4718–4723
- Smith, B.O. *et al.* (2002) Structure of the C3b binding site of CR1 (CD35), the immune adherence receptor. *Cell* 108, 769–780
- Hare, B.J. *et al.* (1999) Structure, specificity and CDR mobility of a class II restricted single-chain T-cell receptor. *Nat. Struct. Biol.* 6, 574–581

- 37 Jasanoff, A. *et al.* (1998) Structure of a trimeric domain of the MHC class II-associated chaperonin and targeting protein II. *EMBO J.* 17, 6812–6818
- 38 Standing, K.G. (2003) Peptide and protein *de novo* sequencing by mass spectrometry. *Curr. Opin. Struct. Biol.* 13, 595–601
- 39 Sobott, F. and Robinson, C.V. (2002) Protein complexes gain momentum. *Curr. Opin. Struct. Biol.* 12, 729–734
- 40 Fenn, J.B. (2003) Electrospray wings for molecular elephants (Nobel lecture). *Angew Chem-Int Ed. Engl.* 42, 3871–3894
- 41 Tanaka, K. (2003) The origin of macromolecule ionization by laser irradiation (Nobel lecture). *Angew. Chem.-Int. Ed. Engl.* 42, 3860–3870
- 42 Standing, K.G. (2000) Timing the flight of biomolecules: a personal perspective. *Int. J. Mass Spectrom.* 200, 597–610
- 43 Cristoni, S. and Bernardi, L.R. (2003) Development of new methodologies for the mass spectrometry study of bioorganic macromolecules. *Mass Spectrom. Rev.* 22, 369–406
- 44 Michalski, W.P. and Shiell, B.J. (1999) Strategies for analysis of electrophoretically separated proteins and peptides. *Anal. Chim. Acta* 383, 27–46
- 45 Hoofnagle, A.N. *et al.* (2003) Protein analysis by hydrogen exchange mass spectrometry. *Annu. Rev. Biophys. Biomol. Struct.* 32, 1–25
- 46 Kaltashov, I.A. and Eyles, S.J. (2002) Studies of biomolecular conformations and conformational dynamics by mass spectrometry. *Mass Spectrom. Rev.* 21, 37–71
- 47 Stoeckli, M. *et al.* (2001) Imaging mass spectrometry: A new technology for the analysis of protein expression in mammalian tissues. *Nat. Med.* 7, 493–496
- 48 Nedelkov, D. and Nelson, R.W. (2003) Surface plasmon resonance mass spectrometry: recent progress and outlooks. *Trends Biotechnol.* 21, 301–305
- 49 Tito, M.A. *et al.* (2001) Probing molecular interactions in intact antibody–antigen complexes, an electrospray time-of-flight mass spectrometry approach. *Biophys. J.* 81, 3503–3509
- 50 Parker, C.E. and Tomer, K.B. (2002) MALDI/MS-based epitope mapping of antigens bound to immobilized antibodies. *Mol. Biotechnol.* 20, 49–62
- 51 Hinshelwood, J. *et al.* (1999) Identification of the C3b binding site in a recombinant vWF-A domain of complement factor B by surface-enhanced laser desorption-ionisation affinity mass spectrometry and homology modelling: implications for the activity of factor B. *J. Mol. Biol.* 294, 587–599
- 52 Hofsteenge, J. *et al.* (1999) The four terminal components of the complement system are C-mannosylated on multiple tryptophan residues. *J. Biol. Chem.* 274, 32786–32794
- 53 Lemmel, C. *et al.* (2004) Differential quantitative analysis of MHC ligands by mass spectrometry using stable isotope labeling. *Nat. Biotechnol.* 22, 450–454
- 54 Flad, T. *et al.* (2003) Development of an MHC-class I peptide selection assay combining nanoparticle technology and matrix-assisted laser desorption/ionisation mass spectrometry. *J. Immunol. Methods* 283, 205–213
- 55 Downard, K.M. (2000) Contributions of mass spectrometry to structural immunology. *J. Mass Spectrom.* 35, 493–503
- 56 Mann, M. *et al.* (2001) Analysis of proteins and proteomes by mass spectrometry. *Annu. Rev. Biochem.* 70, 437–473
- 57 Zhu, H. *et al.* (2003) Proteomics. *Annu. Rev. Biochem.* 72, 783–812
- 58 Weber, P.C. and Salemme, F.R. (2003) Applications of calorimetric methods to drug discovery and the study of protein interactions. *Curr. Opin. Struct. Biol.* 13, 115–121
- 59 Leavitt, S. and Freire, E. (2001) Direct measurement of protein binding energetics by isothermal titration calorimetry. *Curr. Opin. Struct. Biol.* 11, 560–566
- 60 Pierce, M.M. *et al.* (1999) Isothermal titration calorimetry of protein–protein interactions. *Methods* 19, 213–221
- 61 Jelezarov, I. and Bosshard, H.R. (1999) Isothermal titration calorimetry and differential scanning calorimetry as complementary tools to investigate the energetics of biomolecular recognition. *J. Mol. Recognit.* 12, 3–18
- 62 Doyle, M.L. (1997) Characterization of binding interactions by isothermal titration calorimetry. *Curr. Opin. Biotechnol.* 8, 31–35
- 63 Simon, S.L. (2001) Temperature-modulated differential scanning calorimetry: theory and application. *Thermochimica Acta* 374, 55–71
- 64 Yokota, A. *et al.* (2003) The role of hydrogen bonding via interfacial water molecules in antigen–antibody complexation – The HyHEL–10–HEL interaction. *J. Biol. Chem.* 278, 5410–5418
- 65 Sundberg, E.J. *et al.* (2000) Estimation of the hydrophobic effect in an antigen–antibody protein–protein interface. *Biochemistry* 39, 15375–15387
- 66 Kirkitadze, M.D. *et al.* (1999) Central modules of the vaccinia virus complement control protein are not in extensive contact. *Biochem. J.* 344, 167–175
- 67 Willcox, B.E. *et al.* (1999) TCR binding to peptide–MHC stabilizes a flexible recognition interface. *Immunity* 10, 357–365
- 68 Saito, K. *et al.* (2003) Thermodynamic analysis of the increased stability of major histocompatibility complex class II molecule I-E^k complexed with an antigenic peptide at an acidic pH. *J. Biol. Chem.* 278, 14732–14738
- 69 Brooks, C.L. and Case, D.A. (2003) Theory and simulation – The control and timescale of structure and reactivity in biological systems: from peptide folding to cellular networks. *Curr. Opin. Struct. Biol.* 13, 143–145
- 70 McCammon, A.J. and Wolynes, P.G. (2002) Theory and simulation – Enlarging the landscape. *Curr. Opin. Struct. Biol.* 12, 143–145
- 71 Koehl, P. and Levitt, M. (1999) Theory and simulation – Can theory challenge experiment? *Curr. Opin. Struct. Biol.* 9, 155–156
- 72 Karplus, M. and McCammon, J.A. (2002) Molecular dynamics simulations of biomolecules. *Nat. Struct. Biol.* 9, 646–652
- 73 Hansson, T. *et al.* (2002) Molecular dynamics simulations. *Curr. Opin. Struct. Biol.* 12, 190–196
- 74 Honig, B. and Nicholls, A. (1995) Classical electrostatics in biology and chemistry. *Science* 268, 1144–1149
- 75 Elcock, A.H. *et al.* (2001) Computer simulation of protein–protein interactions. *J. Phys. Chem. B* 105, 1504–1518
- 76 Brooijmans, N. and Kuntz, I.D. (2003) Molecular recognition and docking algorithms. *Annu. Rev. Biophys. Biomol. Struct.* 32, 335–373
- 77 Baker, N.A. *et al.* (2001) Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc. Natl. Acad. Sci. U. S. A.* 98, 10037–10041
- 78 McDonald, S.M. *et al.* (1995) Determination of the pK(a) values of titratable groups of an antigen–antibody complex, HyHEL-5-hen egg lysozyme. *Protein Eng.* 8, 915–924
- 79 Norel, R. *et al.* (2001) Electrostatic contributions to protein–protein interactions: fast energetic filters for docking and their physical basis. *Protein Sci.* 10, 2147–2161
- 80 Mallik, B. *et al.* (2003) Conformational interconversion in compstatin probed with molecular dynamics simulations. *Proteins* 53, 130–141
- 81 Morikis, D. and Lambris, J.D. (2004) The electrostatic nature of C3d–complement receptor 2 association. *J. Immunol.* 172, 7537–7547
- 82 Froloff, N. *et al.* (1997) On the calculation of binding free energies using continuum methods: application to MHC class I protein–peptide interactions. *Protein Sci.* 6, 1293–1301
- 83 Michielin, O. and Karplus, M. (2002) Binding free energy differences in a TCR–peptide–MHC complex induced by a peptide mutation: a simulation analysis. *J. Mol. Biol.* 324, 547–569