

Glossary

Bold terms within definitions denote other Glossary terms. B = biological meaning or definition, C = crystallographic context, M = mathematical context. Numbers in parentheses next to the entry refer to defining equations.

A

absent reflections – see **systematic absences**

absolute configuration

Absolute configuration is defined by the sequence of the ligands around a chiral center (S = *sinister*, left; R = *rectus*, right) following the **Cahn–Ingold–Prelog priority rules**. The absolute configuration of a molecule can be determined by **anomalous X-ray diffraction** experiments.

absolute scale – see **Wilson scaling**

absorption edge – see **X-ray absorption edge**

absorption edge scan – see **X-ray excitation scan**

accuracy

A measure of how close an experimental mean value or **expectation value** is to the true value. *Compare* **precision**.

acentric molecule

Molecule that possesses no center of **inversion**.

acentric structure factor (or reflection)

General structure factor that lacks **centrosymmetry** and has no **phase restricted reflections**.

acetyl (–COCH₃)

Chemical group derived from acetic acid (CH₃COOH). Acetyl groups are important in metabolism and are also added covalently to some proteins as a **posttranslational modification**.

acetylcholine (ACh)

Neurotransmitter that functions at cholinergic synapses. Found in the brain and peripheral nervous system. The neurotransmitter at vertebrate neuromuscular junctions.

achiral

Not **chiral**

acid

A proton donor. Substance that releases **protons** (H⁺) when dissolved in water, forming **hydronium ions** (H₃O⁺) and lowering the pH.

active site

Region of an enzyme surface to which a substrate molecule binds in order to undergo a catalyzed reaction.

acyl group (–CO–R)

Functional group derived from a carboxylic acid.

additive (C)

An ion or molecule, added to the crystallization cocktail aiming to alter or improve protein crystallization behavior.

ADME, ADME/T

Acronym in pharmacology for **absorption**, **distribution**, **metabolism**, **excretion** (and **toxicity**), which describes the disposition of a pharmaceutical compound within an organism.

affinity

The strength of binding of a molecule to its ligand at a single binding site.

affinity capture

Enrichment of a protein construct with an **affinity tag** by batch methods or **affinity chromatography**.

affinity chromatography

Type of chromatography in which the protein mixture to be purified is passed over an affinity matrix to which specific ligands for the required protein are attached, so that the protein is retained on the matrix.

affinity constant (association constant) (K_a)

Measure of the strength of binding of the components in a complex. For components A and B and a binding equilibrium $A + B \leftrightarrow AB$, the association constant is given by $[AB]/[A][B]$; the tighter the binding between A and B, the larger the association constant. K_a is the inverse of the **dissociation constant** K_d .

affinity tag

Small **fusion peptide**, linked to a **protein construct**, causing it to bind to an affinity matrix. *See also* **affinity chromatography**.

agonic (C) – see **triclinic**

alanine

Small L- α -amino acid with –CH₃ side chain, somewhat hydrophobic.

alkyl group (C_nH_{2n+1})

General term for a group of covalently linked carbon and hydrogen atoms such as methyl (–CH₃) or ethyl (–CH₂CH₃) groups. Usually exist as part of larger organic molecules. On their own they form extremely reactive free radicals.

allele

One of several alternative forms of a gene. In a diploid cell each gene will typically have two alleles, occupying the corresponding position (locus) on homologous chromosomes.

allostery (adjective **allosteric**)

Change in a protein's conformation brought about by the binding of a regulatory ligand (at a site other than the protein's **active site**), or by covalent modification. The change in **conformation** alters the activity of the protein and can form the basis of directed movement.

alpha helix (α -helix)

Common folding pattern in proteins, in which a linear sequence of amino acids folds into a right-handed helix stabilized by internal hydrogen bonding between backbone atoms. Figure 2-11.

ambiguity – see **substructure handedness, phase ambiguity**

amide

Molecule containing a **carbonyl group** linked to an **amine**.

amine

Chemical group containing nitrogen and hydrogen, positively charged at neutral pH.

amino acid

Organic molecule containing both an **amino group** and a **carboxyl group**. **Proteinogenic amino acids** serving as building blocks of proteins are L- α -amino acids, where the **amino** and **carboxyl** groups linked to the same C α carbon atom in 2S **absolute configuration**. *See also* **CORN rule**.

amino group (–NH₂)

Weakly basic functional group derived from ammonia (NH₃) in which one or more hydrogen atoms are replaced by another atom. In aqueous solution the amino group can accept a proton and carry a positive charge (–NH₃⁺).

amino terminus

First amino acid residue in a polypeptide chain, with free NH₂ group, generally **protonated** at physiological pH, or decorated with posttranslationally modified amino group.

amorphous (C)

Non-crystalline, vitreous, of no defined shape.

amphipathic

Having both hydrophobic and hydrophilic regions, as in a **phospholipid** or a **detergent** molecule.

amphiphile

Small **detergent**-like molecule used in **membrane protein** crystallization.

amplitude

In a transversal wave, the magnitude at the peak of the wave. In the case of X-rays, the maximum of the electromagnetic field vector. *See also* **structure factor amplitude**.

angstrom (Å)

Unit of length, non-SI, but conveniently of the same order of magnitude as atom size. $1 \text{ Å} = 10^{-8} \text{ cm} = 10^{-10} \text{ m} = 0.1 \text{ nm}$.

anion

Negatively charged ion.

anisotropic

Different physical properties in different directions, such as refractive index, different **diffraction limits**, or **anisotropic displacement parameters** of atoms.

anisotropic B-factor – *see* **anisotropic displacement parameters****anisotropic displacement ellipsoid**

Graphical representation of the anisotropic atomic displacement in the form of an ellipsoid with principal axes the diagonal elements of the **anisotropic displacement tensor**.

anisotropic displacement parameter refinement

Also known as anisotropic *B*-factor refinement, requires six adjustable parameters instead of one in **isotropic B-factor [Q1]** refinement. Due to increased number of parameters, generally only possible at resolution of $\sim 1.4 \text{ Å}$ or higher. Can be accomplished in groups by **TLS refinement** at lower resolution.

anisotropic displacement parameters

Define the anisotropic displacement of an atom from its mean position, expressed as the six parameters defining the **anisotropic displacement tensor**.

anisotropic displacement tensor

Symmetric matrix $\mathbf{U} = \langle \mathbf{u}\mathbf{u}^T \rangle$, defined by the mean **dyad product** of the displacement vector \mathbf{u} . Can be graphically represented as an anisotropy ellipsoid, with its principal axes given by the diagonal elements of \mathbf{U} .

anisotropic overall B-factor

Applied during scaling of anisotropic diffraction data, that is, an overall scattering attenuation ellipsoid instead of an **isotropic overall B-factor**.

anisotropic scaling

Applied to account for anisotropic diffraction limits in the form of **anisotropic overall B-factor**. In extreme cases, combined with anisotropic data truncation.

anisotropic scattering tensor

In the presence of polarization anisotropy as used in anisotropic anomalous scattering, the scalar anomalous scattering factors become tensors.

anisotropy tensor

The **dyad product** of an anisotropic property vector, for example **anisotropic displacement tensor U**.

annealing

A process of relaxing stress or high energy states in a material by temperature cycling. Also used to describe minimization procedure used in crystallographic **refinement**.

anomalous difference Patterson map

Patterson map formed from $(F_{PA}^+ - F_{PA}^-)^2$ as Fourier coefficients.

anomalous differences

The anomalous intensity differences $\Delta I = I^+ - I^-$ or differences in

structure factor amplitudes $\Delta F = F^+ - F^-$ between members of a **Bijvoet pair**.

anomalous diffraction

Change of intensities between certain related reflections (**Bijvoet pairs**) due to presence of **anomalous scattering** contributions, caused by anomalous X-ray dispersion.

anomalous diffraction phasing – *see* **anomalous phasing****anomalous diffraction ratio** (6-64)**anomalous dispersion** – *see* **anomalous scattering****anomalous phasing**

Experimental phasing methods that exploit differences in intensities between certain related reflections (**Bijvoet pairs**) due to presence of **anomalous scattering** contributions resulting from phasing marker atoms. *See also* **SAD, MAD, S-SAD**.

anomalous scattering

In the vicinity of an absorption edge the **atomic scattering factor** becomes wavelength- or energy-dependent (anomalous dispersion). The anomalous contributions are a real part f' and an imaginary part $i \cdot f''$ rendering the scattering factor complex. The complex contributions change both phase and amplitude of a general acentric structure factor, leading to a breakdown of **Friedel's law** and generating **anomalous differences** between the members of acentric **Friedel** and **Bijvoet pairs**.

anomalous scattering factor (6-56)

Wavelength-dependent anomalous contributions with real part f' and an imaginary part $i \cdot f''$ rendering the atomic scattering factor complex, thereby introducing additional phase shift and anomalous intensity differences in Bijvoet pairs.

anomalous signal

Colloquial term for **anomalous scattering** contributions in anomalous (difference) diffraction data.

antibiotic

Substance such as penicillin or streptomycin that is toxic to microorganisms. Often a product of a particular microorganism or plant. Used as selective marker in **cloning**.

antibody (immunoglobulin, Ig)

Protein produced by B cells in response to a foreign molecule or invading microorganism. Binds tightly to the foreign molecule or cell, inactivating it or marking it for destruction by phagocytosis or complement-induced lysis. Molecule composed of a heavy and light chain, forming an F_c and two F_{ab} fragments upon enzymatic digestion. Used in crystallization for **antibody scaffolding**.

antibody scaffolding

Utilization of **antibody** fragments for co-crystallization of (membrane) proteins in order to mediate a 3-dimensional network of intermolecular interactions ("scaffold") leading to the formation of a crystal.

anti-freeze – *see* **cryoprotectant****antigen**

A molecule that can induce an adaptive immune response or that can bind to an **antibody** or T-cell receptor.

apo-protein, apo-enzyme

Protein not bound to its ligands or complex partners.

archaea

Archaea or archaeobacteria are single-celled organisms without a nucleus, superficially similar to **bacteria**. At a molecular level, more closely related to bacteria in metabolic machinery, but more similar to **eukaryotes** in genetic machinery. Archaea and bacteria together make up the **prokaryotes**.

Argand diagram

Graphical representation of **complex numbers** in a 2-dimensional diagram, with the real axis horizontal and the imaginary axis vertical.

arginine

Positively charged L- α -amino acid, side chain $-(\text{CH}_2)_3\text{NHC}(\text{NH}_2)_2^+$ with planar conjugated guanidinium group.

ARP/wARP

Dummy atom refinement and model building program, available on the web.

asparagine

Polar L- α -amino acid with $-\text{CH}_2\text{CONH}_2$ side chain. Hydrogen bond donor.

aspartate, aspartic acid

Charged, acidic L- α -amino acid with $-\text{CH}_2\text{COO}^-$ side chain. Hydrogen bond acceptor.

association constant – see **affinity constant****associativity (M)**

Associativity implies that $a(bc) = (ab)c$ (which does not automatically apply **commutativity**, that is, $ab \neq ba$).

asymmetric

Not possessing any symmetry, **chiral**.

asymmetric unit (cell)

Smallest part of a crystal structure, containing the **motif**, from which upon application of **space group symmetry operations** the complete unit cell of the crystal is generated.

atomic coordinates (C)

Define the positions of the atoms in a crystal structure in reference to a common origin. Can be in **Cartesian coordinate** form (world coordinates) or **crystallographic** or **fractional coordinates**.

atomic displacement parameter

Also called *B*-factor, a measure related to the atomic displacement (vector) of an atom from its mean position. The nature of atomic displacement is variable; can be thermal vibration about the equilibrium position, static disorder of atomic position in different unit cells, or dynamic effects of group movement. Atomic displacement can be accounted for in **isotropic** parameterizations (overall or individual refinement) or **anisotropic** parameterization (individual or group **TLS refinement**).

atomic force microscope (AFM)

The atomic force microscope is a high-resolution type of scanning probe microscope, with a resolution of fractions of a nanometer, more than 1000 times better than the optical diffraction limit for a light microscope.

atomic resolution (C)

Diffraction limit beyond which individual atoms can be discerned in electron density, commonly meaning better than $\sim 1.2 \text{ \AA}$.

atomic scattering factor, *f*

Measure for the X-ray scattering power of an atom, equal to the number of **electrons** at zero scattering angle, and decreasing rapidly with scattering angle. The atomic scattering factor is the **Fourier transform** of the atomic electron density, commonly approximated by a spherically symmetric superposition of **Gaussian** functions. Most common is the approximation by nine **Cromer–Mann coefficients**. Only at ultra-high **resolution** (0.8 \AA and better) the anisotropy of the electron density due to valence electrons (charge density deformation) needs to be considered. The units of the scattering factor are electrons.

ATP (adenosine 5'-triphosphate)

Nucleoside triphosphate composed of adenine, ribose, and three phosphate groups. The principal carrier of chemical energy in cells. The terminal phosphate groups are highly reactive in the sense that their hydrolysis, or transfer to another molecule, takes place with the release of a large amount of free energy.

autoconvolution – see **autocorrelation****autocorrelation (9-25)**

Convolution of the same function with its inverse; for example, the **Patterson function** is an autocorrelation of the electron density, $P(\mathbf{u}) = \rho(\mathbf{r}) \otimes \rho(-\mathbf{r})$

autoinduction

Elegant method to induce expression in diauxic *E. coli* when switching from glucose to lactose metabolism.

Avogadro's number (Loschmidt's number)

Number of molecules in a mole of material, $N_A = 6.022142 \times 10^{23} \text{ mol}^{-1}$.

B**B-factor** – see **atomic displacement parameter****B-factor distribution**

A positively skewed distribution that is specific for each molecule and can indicate abnormal *B*-factor restraints or scaling.

B-factor restraints

Restraints placed on the difference in displacement parameters (*B*-factors) of adjacent atoms.

B-factor scaling – see **Wilson scaling****Babinet's principle**

States that the diffraction pattern of an object is the same as that of a hole of the same size and shape as the object, but with opposite phase.

backbone atoms – see **main chain atoms****bacteriophage (phage)**

Any virus that infects **bacteria**. Phages were the first organisms used to analyze the molecular basis of genetics, and are now widely used as cloning vectors.

bacteriorhodopsin

Pigmented protein found in the plasma membrane of a salt-loving archaean, *Halobacterium salinarium* (*Halobacterium halobium*). Pumps protons out of the cell in response to light. Among the first **integral membrane proteins** whose 3-dimensional structure was determined. Figure 3-45.

bacterium (plural bacteria) (eubacterium)

Member of the domain bacteria, one of the three main branches of the tree of life (archaea, bacteria, and **eukaryotes**). Bacteria and archaea both lack a distinct nuclear compartment, and together comprise the **prokaryotes**.

baculovirus

Virus that exclusively infects insect cells; used as corresponding expression vector.

ball and stick model

A model of a molecular structure in which the covalent bonds are represented as sticks, with the atoms at the intersection of the bond sticks represented as balls. *Compare* **stick model**.

bandwidth

Generally the region between a lower and upper frequency or energy cut-off selected from a continuous spectrum. In spectroscopy also the **full width at half maximum** of an emission line.

base

A substance that can reduce the number of **protons** in solution, either by accepting H^+ ions directly, or by releasing OH^- ions, which then combine with H^+ to form H_2O . The purines and pyrimidines in DNA and RNA are organic nitrogenous bases and are often referred to simply as bases.

base excision repair

DNA repair pathway in which single faulty bases are removed from the DNA helix and replaced. *Compare* **nucleotide excision repair**.

base pair (Watson–Crick base pair)

Two nucleotides in an RNA or DNA molecule that are held together by hydrogen bonds—for example, G paired with C, and A paired with T or U.

basic (alkaline)

Having the properties of a base.

basis (C)

The *n* linearly independent vectors with a common origin that span an *n*-dimensional space.

batch crystallization

Direct combination of protein solution and crystallization cocktail. Often conducted under paraffin oil without occurrence of significant **vapor diffusion**, or under a mixture of paraffin oil and silicone oil which is semi-permeable to water.

Bayes factor (12-104)

Confidence measure expressing the relative merits of competing hypotheses, similar to **likelihood ratios**, but with consideration of prior (model) odds. Sidebar 12-10.

Bayes' theorem (7-17)

Law of formal inductive inference: the **posterior probability** of a model or hypothesis is proportional to its **data likelihood function** times its **prior probability**.

Bayesian inference (reasoning)

Form of inductive inference, where prior probabilities are explicitly used in the formulation of joint probabilities and likelihood functions.

Bayesian posterior – see **posterior probability****bending magnet**

Located between the straight sections of a synchrotron storage ring, strong bending (electro) magnets are used to deflect the electrons and steer them on a closed circular path. At the same time, due to the acceleration they experience, the electrons emit polychromatic synchrotron radiation, used for X-ray diffraction experiments. Figure 8-6.

Bessel functions

Set of ubiquitous functions obtained as solutions of the Bessel differential equation. Modified Bessel functions of the first kind $I_n(x)$ of order n are exponentially increasing functions obtained during the marginalization of nuisance variables from complex probability distribution functions.

beta-bulge

Distortion caused in β -sheets by insertion of an extra residue, perturbing the regular hydrogen bonding pattern.

beta-hairpin turn

Short secondary structure element with distinct n to $n + 3$ hydrogen bonding pattern connecting two antiparallel β -strands.

beta-sheet (β -sheet)

Common structural motif in proteins in which different strands of the polypeptide chain run either parallel or antiparallel alongside each other, joined together by hydrogen bonding between atoms of the polypeptide backbones.

bias (C) – see **model bias****bias (M)**

Tendency of a result toward a certain outcome, often implying non-randomness in statistical sampling or dependence on hidden parameters.

bidentate

Ligand or additive molecule in crystallization with two functional groups.

bifurcated β -sheet

A β -sheet structure where a strand is shared and connects two sheets in different orientation. Figure 2-16.

bifurcated hydrogen bond

Hydrogen bond in which a hydrogen atom is shared between two hydrogen bond acceptors.

Bijvoet difference

The anomalous intensity differences $\Delta I = I^+ - I^-$ or differences in structure factor amplitudes $\Delta F = F^+ - F^-$ between members of a **Bijvoet pair**.

Bijvoet pair

Each of the m symmetry related equivalents hR_m of reflection h (or F^+) has a **Friedel mate**, $(-h)R_m$ (or F^-). These reflections F^+ and F^- form Bijvoet pairs. Similarly, starting from the Friedel mate $-h$ (or F^-) of a reflection h and then applying the m point group

followed by **inversion** generates $(-h)R_m$ reflections (or F^+), and the pairs F^+ and F^- are again Bijvoet mates. In the general acentric case, **anomalous differences** exist between Bijvoet pairs.

binding pocket

Deep cleft or depression on the surface of a protein molecule that interacts with another molecule through noncovalent bonding.

binding site

Region on the surface of one molecule (usually a protein or nucleic acid) that interacts with another molecule through noncovalent bonding.

birefringence

Splitting of light into two components with different planes of polarization as a result of different refractive index in optically **anisotropic** media, such as protein crystals, where it can be detected with the aid of a polarizing microscope.

bitopic membrane protein

Protein passing through the lipid membrane of a cell with one or more transmembrane helix, connecting an extracellular ectodomain and the intracellular cytosolic domain, which in the case of a single transmembrane helix often can be expressed separately. Compare **monotopic membrane protein**.

body-centered

A lattice or unit cell with centering vector $I = (\frac{1}{2} \frac{1}{2} \frac{1}{2})$. Possible in **tetragonal (tI)**, **orthorhombic (oI)**, and **cubic (cI)** lattices.

Bragg equation

Fundamental relation between lattice spacing $d(hkl)$ of reflecting lattice planes hkl of a crystal structure and the diffraction angle of a X-ray reflection: $d(hkl) = n\lambda / (2 \sin \theta)$.

Bragg reflection

An X-ray diffraction spot or discrete reflection fulfilling the **Bragg equation**. Compare **thermal diffuse scattering**.

Bravais centering – see **Bravais translations****Bravais lattices**

The combination of lattice symmetry with requirements for translational periodicity generates in addition to the six **primitive** lattices additional eight **centered lattice** types, collectively termed the 14 Bravais lattices.

Bravais symbols – see **Bravais translations****Bravais translations**

The translational centering vectors for the eight centered **Bravais lattices** (Figure 5-30). The centering vectors are $(\frac{1}{2} \frac{1}{2} 0)$ for *C*-centering; $(\frac{1}{2} \frac{1}{2} 0)$, $(\frac{1}{2} 0 \frac{1}{2})$, and $(0 \frac{1}{2} \frac{1}{2})$ for *F*-centering; $(\frac{1}{2} \frac{1}{2} \frac{1}{2})$ for *I*-centering; and $(\frac{2}{3} \frac{1}{3} \frac{1}{3})$, $(\frac{1}{3} \frac{2}{3} \frac{2}{3})$ for *R*-centering. Table 6-6.

Bremsstrahlung

Continuous part of the X-ray emission spectrum emanated by a laboratory **X-ray generator**. Figure 8-2.

brilliance

A measure for **photon flux** that takes the beam divergence into account, defined as the number of photons per second that pass through an area of 1 mm² with a given divergence measured in millirad: $B_\lambda = \text{photons} \cdot \text{s}^{-1} \cdot \text{mm}^{-2} \cdot \text{mrad}^{-2}$.

buffer

Solution of weak acid or weak base that maintains (buffers) the pH when small quantities of acid or base are added, or when solution is diluted. See also **crystallization cocktail**.

bulk solvent correction

Disordered solvent attenuates low resolution Bragg reflections, which is accounted for by bulk solvent correction.

C**C-centering**

Lattice centering with (Bravais) centering vector $C = (\frac{1}{2} \frac{1}{2} 0)$. Possible in monoclinic (*mC*) and orthorhombic (*oC*) lattices. Figure 5-30.

C-terminus – see **carboxyl terminus**

Cahn–Ingold–Prelog priority rules

Rules for assignment of ligand priority in establishing the **absolute configuration** of a **chiral center**. In first order, following the size of the ligand, for α -amino acids the sequence is from C α to H, N, C, C β . See also **CORN rule**.

calmodulin

Ubiquitous intracellular Ca²⁺-binding protein that undergoes a large conformation change when it binds Ca²⁺, allowing it to regulate the activity of many target proteins. In its activated (Ca²⁺-bound) form, it is also called Ca²⁺/calmodulin.

calorie

Unit of energy, equal to 4.184 joules (J). One calorie (cal) is the amount of heat needed to raise the temperature of 1 gram of water by 1°C.

cancer

Disease featuring abnormal and improperly controlled cell division resulting in invasive growths, or tumors, that may spread (metastasize) throughout the body.

capsid

Protein coat of a virus, formed by the self-assembly of one or more types of protein subunit into a geometrically regular structure, frequently of icosahedral symmetry. Figure 2-2.

carbohydrate

General term for sugars and related compounds containing carbon, hydrogen, and oxygen, usually with the empirical formula (CH₂O)_n. Structural elements in **glycoproteins**.

carbonyl group (>C=O)

Carbon atom linked to an oxygen atom by a double bond.

carboxyl group (–COOH)

Carbon atom linked both to an oxygen atom by a double bond and to a hydroxyl group. Molecules containing a carboxyl group are weak carboxylic acids.

carboxyl terminus

Last amino acid residue in a polypeptide chain, with free COOH group, **deprotonated** under physiological pH.

Cartesian coordinates (basis)

An **orthogonal** coordinate basis [0, X, Y, Z] in **real space**, with dimensions suitable to the object to be represented (Å for molecules); world coordinates.

Cauchy distribution – see **Lorentz–Cauchy distribution**

CCP4 (Collaborative Computing Project 4)

A major collection of popular crystallographic programs bundled with a graphical user interface (CCP4i) and supported by a large user community.

cDNA

DNA molecule made as a copy of mRNA and therefore lacking the introns that are present in eukaryotic genomic DNA.

cell (B)

The cell is the structural and functional unit of all known living organisms, separated by a membrane from its environment.

cell (C) – see **unit cell**

cell-free system

Fractionated cell homogenate that retains the protein synthesis function of the intact cell.

cell line

Population of cells of plant or animal origin capable of dividing indefinitely in culture.

cell parameters – see **unit cell parameters**

center of mass

Defined as mass-weighted average, of i atoms in a molecule,

$$C = \sum_i \mathbf{r}_i m_i / \sum_i m_i$$

centered lattice

The eight centered **Bravais lattices** are mC , oI , oC , oF , tI , hR , cI , and cF .

central limit theorem

States that the distribution obtained by random sampling from almost any unimodal parent probability distribution will be a **normal distribution**, irrespective of the shape of the parent distribution.

central moment (n th central moment)

The n th moment of a distribution about the mean is defined as $\mu_n = E(x - x_0)^n$ with the **expectation value** E defined as the integral over the distribution. The first central moment is zero, the second moment about x_0 is the **variance** σ_x^2 , the third is the **skewness**, and the fourth is the **kurtosis**, a measure how “squat” or “peaky” a distribution is. Compare **raw moment**.

centric Patterson projections – see **Harker sections**

centric reflections

Centric reflections are **phase restricted reflections** resulting from scattering by centrosymmetrically related atoms, that is, related by the point group symmetry of the crystal structure. They are located in centric zones of reciprocal space, and have phases of ϕ or $\phi + 180^\circ$, often (in the absence of transitional symmetry elements) but not necessarily 0° or 180° . **Bijvoet pairs** of centric reflections cannot carry any **anomalous difference** signal even in the presence of **anomalous scattering** contributions, because they also fulfill the symmetry relations of the corresponding space (point) group. Not all space groups have centric reflections; for example, in $P3$ none of the three point group operators \mathbf{R} applied to \mathbf{h} generates a reflection \mathbf{hR} that is also a Friedel mate $-\mathbf{h}$ of itself.

centroid phases (10-43)

Phases defined by the centroid or weighted average of the possible phase probabilities, obtained by integration over all phase angles.

centrosymmetry

Centrosymmetry relates any object with positional coordinate vector \mathbf{x} to an inverted object with coordinate vector $-\mathbf{x}$. Centrosymmetry (**inversion**) changes handedness of the motif and is therefore not possible in native protein crystal structures.

change of handedness – see **substructure handedness**

channel (membrane channel)

Transmembrane protein complex that allows inorganic ions or other small molecules to diffuse passively across the lipid bilayer **membrane**.

chaotropic

A chaotropic reagent (chaotrope) is a substance which disrupts the 3-dimensional structure in proteins, DNA, or RNA and denatures them. Chaotropic agents interfere with stabilizing intramolecular interactions mediated by noncovalent forces such as hydrogen bonds, van der Waals forces, and hydrophobic effects.

chaperone (molecular chaperone)

Protein that facilitates the proper folding of other proteins, or helps them avoid misfolding, for example heat shock proteins (Hsp).

characteristic radiation – see **X-ray fluorescence**

charge, charged

The unit of charge is the charge of one electron, $1.602176487(40) \times 10^{-19}$ Coulomb (Cb). Charged **residues** carry a positive or negative charge and are involved in charged or ionic interactions, also called salt bridges.

charge coupled device (CCD)

Imaging device built from a semiconductor array that transforms radiation into electric signal.

charge flipping

Substructure solution method related to **solvent flipping**, where beyond a certain threshold value the low electron density is

subtracted from the high electron density, new structure factors are calculated, and the process is cycled.

chemical shift

Term used in spectroscopic techniques for energy shift of absorption signal (X-ray absorption) or resonance signal (NMR) due to a specific local chemical environment.

Cheshire cell

The limits of the Cheshire cell determine the space that has to be searched during translational searches.

Cheshire group

The different possible origin and **enantiomorph** choices combined with the symmetry operators of a space group generate the complete set of operators of the so called Cheshire groups. Cheshire groups can be visualized as the symmetry groups of the symmetry operators, and play a role in crystallographic search procedures. They are formally derived as the Euclidean normalizers of the space groups.

chi-square, χ^2 (7-35)

Sum of normalized residuals squared, used in **refinement** target functions and statistical χ^2 tests.

chi-square test (7-36)

The reduced χ^2 is employed in the statistical χ^2 tests deciding whether a distribution is a **normal distribution** ($\chi^2 = 1$). Used to adjust error models, for example during data processing, to reflect a normal distribution.

chiral center, asymmetric center

Atom (generally sp^3 hybridized carbon) with four different, non-planar ligands. Cannot undergo **inversion** or **mirror operation** without changing **handedness**.

chiral, chirality

A molecule that does not possess inversion or mirror symmetry is chiral. In proteins chirality results from the presence of **asymmetric** chiral $C\alpha$ carbon atoms with four different covalently bound groups. **Isoleucine** and **threonine** have a second chiral center at $C\beta$; their **absolute configuration** is (2S, 3R). Chirality is associated with optical activity.

chiral space group

Space groups that can act on a **chiral** motif, that is, contain no **mirror operations** or **inversions**. Sixty-five of the 230 space groups are chiral (Table 6-6).

chiral volume

A scalar triple vector product $\mathbf{A} \cdot (\mathbf{B} \times \mathbf{C})$ originating at the central atom with the vector order following the **Cahn-Ingold-Prelog priority rules**. For L-amino acids with absolute (2S) configuration, the sign of the chiral volume is invariably positive, and computes to about 2.5 \AA^3 .

chromatography

Broad class of biochemical techniques in which a mixture of substances is separated by charge, size, hydrophobicity, noncovalent binding affinities, or some other property by allowing the mixture to partition between a moving phase and a stationary phase.

circular dichroism (spectroscopy)

Circular dichroism is observed when optically active matter absorbs left and right hand circular polarized light slightly differently. Used to experimentally determine secondary structure composition of a protein.

cis

On the same or near side. Compare *trans*.

clathrate

A local cage of ordered water molecules, similar in structure to that of ice, around a molecule such as a protein, whose formation leads to a loss of **entropy**.

clone, cloning

Population of identical individuals (cells or organisms) formed by repeated (asexual) division from a common ancestor. Creation of multiple copies of a gene by growing a clone of carrier cells (such

as *E. coli*) into which the gene has been introduced, and from which it can be recovered, by **recombinant DNA** techniques.

cloning vector

Small DNA molecule, usually a circular plasmid, which is used to carry the fragment of DNA to be cloned into the recipient cell, and which enables the DNA fragment to be replicated.

close-packed lattice

The face-centered cubic lattice (*cF*) and hexagonal close-packed lattice (*hP*) achieve this same highest possible packing density for spheres with a coordination number of 12 and a packing density of ~74%.

closure requirement (M)

Requirement that operations on members of a set forming a mathematical **group** under these operations are again group members.

closure residual (C)

The difference between observed and calculated **heavy atom derivative** structure factor amplitudes, $F_{\text{PH}}(\text{obs}) - F_{\text{PH}}(\text{calc})$.

CNS (C)

A versatile high-level programming language and program package for macromolecular refinement (Crystallography and NMR System) by Axel Brünger and colleagues. Unique feature **simulated annealing torsion angle refinement**. *CNS* also includes all aspects of phasing (including MAD, SAD, SIR/MIR) and density modification. **Cross-validation** is used throughout the program.

co-crystallization

Crystallization of a complex of a protein and a binding partner from an incubated solution of both components. *Compare* **soaking**.

codon

Sequence of three nucleotides in a DNA or mRNA molecule that represents the instruction for incorporation of a specific amino acid into a growing polypeptide chain.

coenzyme

Small molecule tightly associated with an enzyme that participates in the reaction that the enzyme catalyzes, often by forming a covalent bond to the substrate. Examples include biotin, PLP, NAD, and coenzyme A.

cofactor

Inorganic ion or coenzyme required for an enzyme's activity.

coherence

Constant phase relation in two or more waves or photon wave packets. Normal light sources and X-rays are incoherent sources of radiation; lasers are coherent radiation sources.

coherence length (C)

Distance over which the electromagnetic scattering process is coherent.

coherent scattering

Scattering process during which defined phase relations are maintained.

coiled-coil

Stable rod-like protein structure formed by two or more α -helices coiled around each other.

collimator

Device that passively limits divergence of X-rays through entrance and exit slits.

colony PCR

Direct screening of bacterial colonies by **polymerase chain reaction** to screen for correct DNA vector constructs.

commensurately modulated structure

Crystal structure in which motifs vary in a systematic fashion compatible with the periodic translation period of the original lattice. Occurs in the form of **translational NCS** and manifests itself in Patterson space as well as in reciprocal space as a superstructure pattern. *Compare* **incommensurately modulated structure**.

commutativity (M)

Commutativity implies that the outcome of an operation is independent of the order of elements, that is, $ab = ba$.

competent cell (B)

Competence is the ability of a cell to take up extracellular DNA from its environment.

complex (B)

Molecules bound to each other, commonly a protein and ion, small molecule, ligand, DNA or other protein molecule, most frequently formed by noncovalent interactions.

complex conjugate (M)

The complex conjugate of a **complex number** is obtained by changing the sign of its imaginary part; in the case of a complex **structure factor** by changing the sign of the **phase angle**.

complex number

A complex number or function has a real part and an imaginary part, with the imaginary unit i defined as the square root of -1 . They are conveniently represented in an **Argand diagram**.

complex structure factor

The resultant electromagnetic wave describing a diffracted X-ray, expressed as a complex number.

composite omit map – see omit map**Compton scattering**

Inelastic scattering process involving electrons and electromagnetic radiation, generally leading to loss of energy of scattered radiation or particle.

conditional probability

A probability or probability function that depends on certain conditioning parameters or **conditioning information**. For example, the probability of the outcome of a certain face of a die depends on the number of faces the die has.

conditioning information – see conditional probability**conditions limiting reflections – see systematic absences****confidence interval (7-10)**

Region or range of a probability function within which an observation will fall with a given probability, expressed in standard deviations or percent probability.

conformation

The spatial arrangement of atoms in a molecule with respect to bonds, angles, and torsions. *See also* **absolute configuration**.

consensus sequence

Average or most typical form of a sequence that is reproduced with minor variations in a group of related DNA, RNA, or protein sequences. Indicates the nucleotide or amino acid most often found at each position. Preservation of a sequence generally implies that it is functionally important.

constraints

Mathematical relation between two or more parameters in a multivariate model.

contours – see electron density maps**convergence (C)**

Convergence is approached or reached when, in an iterative procedure such as **macromolecular refinement**, the refinement parameter shifts become negligible.

convergence radius

Measure for parameter deviation from which a global minimum can be reached by a given optimization method. About 1.7 Å coordinate r.m.s.d. from target for torsion angle parameterization, 1.2 Å for coordinate refinement.

convolution (9-24)

Operation involving two functions, defined by a **convolution integral** and similar to probing or point-wise multiplying one function with the other.

convolution integral (9-24)**convolution theorem – see Fourier convolution theorem****coordinate systems**

Spanned by a set of **basis** vectors; most commonly **crystallographic coordinate** systems with **fractional coordinates** and the **Cartesian coordinate** system with coordinates in Å. Can be transformed into each other by affine similarity transformation using the **orthogonalization matrix** and its inverse, the deorthogonalization matrix.

CORN rule

The CORN rule is a practical aid for determining the configuration of **chiral** $C\alpha$ centers of **amino acids**. When the central $C\alpha$ atom is viewed with the H atom pointing toward the observer (or out of the paper plane), the ligand sequence reads “CO-R-N” in a clockwise rotation. Figure 2-5.

correlation coefficient (7-53)

A scale-factor-independent measure for the agreement between observed and calculated quantities. *See also* **real space correlation coefficient**.

COSY (correlation spectroscopy)

2-dimensional homonuclear NMR technique indicating which hydrogen atoms are spin-coupled through bonds. *Compare* NOESY.

covalent bond

Strong, directional chemical bond in which atoms share electrons.

covariance (matrix)

Pair-wise measure for how one parameter or variable in a multivariate function depends on another parameter. The covariance matrix contains the variance of each single parameter in its diagonal.

critical energy, of synchrotron storage ring

Mode of the energy spectrum emitted from a synchrotron (8-1).

Cromer–Mann coefficients (6-15)

The coefficients in the commonly used isotropic nine-parameter Cromer–Mann approximation of the **atomic scattering factor**.

cross-crystal map averaging

Used to improve electron density maps by averaging and density modification of density from the same molecule in different crystal forms.

cross-Fourier map (10-16)

Difference map based on difference amplitudes of native and second derivative and SIR phases of first derivative, used to maintain origin choice between substructures.

cross-Patterson vector

Interatomic distance vector between atoms of different molecules in a crystal structure. Generally longer than **self-Patterson vectors** between atoms in the same molecule.

cross-product – see vector product**cross-rotation – see Patterson cross-rotation****cross-section**

Quantifies the probability of a certain particle–particle or particle–wave interaction such as X-ray or neutron scattering or absorption.

cross-validation

Statistical method to avoid overparameterization of a model by excluding a subset of data from refinement against which a cross-validation statistic such as **R-free** is computed.

cross-validation R-value, R-free (12-35)

R-value calculated for the cross-validation data excluded from refinement.

cryobuffer

A suitable buffered solution of a **cryoprotectant**.

cryocrystallography

Method of collecting diffraction data at **cryogenic** temperatures, largely to minimize **radiation damage** in the case of protein

crystals. Generally requires **cryoprotection** of the protein crystals and/or **flash cooling**.

cryogenic

Very low temperature (below 123 K or $-150\text{ }^{\circ}\text{C}$, $-238\text{ }^{\circ}\text{F}$). *See also* **cryoprotection**.

cryoprotectant

A substance or solution that prevents the formation of crystal-line ice, in common use anti-freeze. Ice formation (taking place in **mother liquor** surrounding a harvested protein crystal) almost invariably degrades or destroys the crystal. Substances acting as cryoprotectants are salts, organic precipitants, PEGs, glycerol, sucrose solutions, and others.

cryoprotection

Prevention of ice formation in the **mother liquor** surrounding harvested protein crystals during cooling, because ice formation almost always degrades or destroys protein crystals. The crystals are cryoprotected through one or more of (i) significant amounts of **cryoprotectants** already in the mother liquor, (ii) swiping or soaking them in cryoprotectant, (iii) **flash cooling** in liquid nitrogen.

crystal

A periodic assembly of identical objects such as atoms or molecules.

crystal class – *see* crystallographic point groups

crystal contacts

Intermolecular contacts that connect symmetry related molecules (including translationally related molecules from neighboring cells) in a **crystal structure**.

crystal faces

Distinct lattice planes forming and defining the macroscopic shape of the crystal.

crystal form

Different **crystal structures** formed by the same material. *Compare* **crystal habit**.

crystal habit

Different appearance of the same **crystal form** of the same material. Shape, isotropy, development of crystal faces, as well as diffraction properties can differ for different crystal habits.

crystal lattice

A periodic construct generated by translation of unit cells along the **lattice vectors**.

crystal structure

A specific arrangement of molecules in a **crystal**, defined by unit cell dimensions, space group, and molecular arrangement. Same **space group** does not mean same crystal structure, even for the same material. *See also* **crystal form**.

crystal system

The seven crystal systems arise from the six **lattice types** under consideration of minimal internal symmetry, which splits the hexagonal lattice into a **trigonal** (only 3-fold internal symmetry) and a **hexagonal** (full 6-fold internal symmetry) system.

crystallization cocktail

Mixture of various **precipitants**, **buffers**, and **additives** used to reduce protein solubility and achieve **supersaturation**.

crystallization diagram

A pseudo-binary **phase diagram** of protein concentration versus precipitant concentration, depicting protein solubility and phase relations of a protein solution together with a depiction of tentative information about kinetic phenomena such as nucleation.

crystallographic coordinates – *see* fractional coordinates

crystallographic fragment screening

A method of experimental ligand screening in which small molecules—fragments of drug-like substances—are soaked or co-crystallized with target molecules. From the binding fragments, drug-like molecules are designed in an iterative process.

crystallographic point group

Point groups are mathematical **groups** of operations that keep the origin fixed (e.g. all symmetry operations intersect at the origin). There is an infinite number of 3-dimensional point groups but only 32 point groups comply with the crystallographic translation restrictions. Only eleven of the crystallographic point groups are enantiomorphic, that is, allowed for **chiral** motifs such as protein molecules. Table 6-6.

crystallographic restriction theorem

Requirement that all lattice points need to superimpose upon each other when a symmetry operation is applied. Follows from properties of the **direction cosine matrix** defining a rotation.

crystallographic rotation axis

An n -fold rotation axis ($n = 2, 3, 4, 6$) in specific orientation (parallel, perpendicular, along space diagonal) to a **unit cell vector**. The orientation of the axes can be found in the **unit cell projections** depicted for each **space group** in the International Tables for Crystallography, Volume A.

crystallographic rotations

Crystallographic rotations are **proper n -fold symmetry operations** limited by crystallographic translational periodicity restrictions to n ($n = 2, 3, 4, 6$) successive rotations by $360/n$ degrees around a **crystallographic rotation axis** in a specific orientation (parallel, perpendicular, along space diagonal) to a unit cell vector.

crystallographic screw axis

An n -fold rotation axis n_m combining n successive **rotations** by $360/n^{\circ}$ parallel to a unit cell vector **k** with m successive translations of km/n along the rotation axis where $n = 2, 3, 4, 6$ and $m = 1$ to $n - 1$. For example, a 4_3 screw axis parallel to unit cell axis **c** generates 3 additional **symmetry equivalent** copies by clockwise rotations by $m \times 90^{\circ}$ plus m translations by cm/n . Figures 5-32, 5-33, 5-34.

cubic

Lattice or crystal structure with cell parameters $a = b = c$, $\alpha = \beta = \gamma = 90^{\circ}$, with primitive (*cP*), *I*-centered (*cI*), and *F*-centered (*cF*) lattice types. Minimal internal symmetry four 3-fold rotation axes along the space diagonals.

cumulative distribution function

Integral over a probability distribution function.

curvature (M)

Second **derivative**, or change of **slope**, of a function. At a maximum of a function the curvature is negative; at the minimum, positive.

cyclic group (C)

A space group that is generated by successions of a single generator, such as a **proper rotation**.

cysteine

Polar L- α -amino acid containing reactive thiol group in side chain $-\text{CH}_2\text{SH}$. Forms disulfide bonds to adjacent cysteines upon oxidation.

cytochrome

Colored heme-containing protein that transfers electrons during respiration, drug metabolism (cytochrome P-450), and photosynthesis.

cytosol

Contents of the main compartment of the cellular cytoplasm, excluding membrane-bounded organelles such as endoplasmic reticulum and mitochondria.

D

dalton

Unified atomic mass unit, Da. One twelfth of the mass of a ^{12}C atom at rest and approximately equal to the mass of a hydrogen atom (1.66054×10^{-24} g). $1\text{ Da} = 1/N_A$ grams where N_A is **Avogadro's number**.

data (C)

Experimental **observations** (structure factor amplitudes or intensities).

data likelihood (function)

Term frequently used for the probability function $\text{prob}(\text{data}|\text{model}, I)$, also known as the sampling probability.

data-to-parameter ratio

A measure for the **determinacy** of a system; with a data-to-parameter ratio > 1 the system is overdetermined, but for sensible parameter refinement, multiple **observations** per adjustable parameter are required.

de Broglie relation

Relates mass m and energy (or **wavelength**) of particles. For particles moving at light speed c , the wavelength is $\lambda = h/mc$ and the frequency $\nu = mc^2/h$ with h the Planck constant. *See also* **energy-wavelength conversion**.

Debye effect

Deviations from linearity in **Wilson plot** for low-resolution data, due to correlation between atomic distances.

Debye-Waller factor (6-16)

The entire scattering vector dependent exponential term bearing the **atomic displacement parameter** or **B-factor** in the (negative) exponent.

deductive inference (logic)

An argument is deductive when the truth of the conclusion follows necessarily as a logical consequence of the premises, that is, its corresponding conditional is a necessary truth. *Compare* **inductive inference**.

denaturation

Dramatic change in conformation of a protein caused by heating or by exposure to chaotropic reagents, usually resulting in the loss of biological function.

density averaging

An averaging procedure whereby electron density of multiple molecules in the asymmetric unit related by **non-crystallographic symmetry** is averaged, generally combined with solvent flipping. Very powerful with high NCS—highly symmetrical virus capsids can be phased starting from a spherical shell as a model.

density modification

Methods for phase improvement and phase extension that utilize the distinctly different properties of protein versus disordered solvent. **Solvent flattening**, **solvent flipping**, and the **sphere of influence algorithm** (often combined with electron **density averaging**) are common density modification procedures.

deorthogonalization matrix

Inverse of the orthogonalization matrix.

deoxyribonucleic acid – see DNA**deoxyribose**

The five-carbon monosaccharide component of DNA. Differs from **ribose** in having H at the C-2 position rather than OH. $C_5H_{10}O_4$.

derivative (C)

A derivative crystal contains (heavy) atoms used as phasing marker atoms.

derivative (M)

A measure of how a mathematical function changes as its parameters change. The first derivative of a function at a given point defines its **slope** (tangent), the second derivative the **curvature** (i.e. the change of the slope).

derivative-difference Fourier map (10-17)

Difference Fourier map use to identify additional low-occupancy sites.

derivative Patterson map

A Patterson map based on coefficients $(F_{PA1} - F_{PA2})^2$ used (historically) to maintain origin choice between derivatives.

detergent

Organic molecules that are amphiphilic, meaning they contain both hydrophobic groups (their “tails”) and hydrophilic groups

(their “heads”). Therefore, they are soluble in both organic solvents and water. They form micelles at concentrations above the critical micelle concentration. Used in crystallization as an additive and to keep (membrane proteins) in solution.

determinacy (limit)

The limiting resolution defined by a data-to-parameter ratio of unity below which a system becomes undetermined. Around 4.5 Å for torsion angle refinement, $\sim 3.5\text{--}3.0$ Å for restrained refinement.

determinant (M)

The determinant of a matrix describing a **linearly independent** set of vectors (or set of equations) is nonzero. In a 3-dimensional **basis** system, right-**handedness** implies a positive determinant; **inversion of handedness** changes the sign of the determinant.

deviate – see residual**dialysis**

Exchange of small molecules, water, or ions across a membrane separating two systems. Process follows the chemical gradient.

dielectric constant

Relative static permittivity, a measure for the electrostatic flux density or polarity of a solvent. For example, water (highly polar) has a dielectric constant of ~ 80 while *n*-hexane (nonpolar) has a dielectric constant of 1.9 at 20°C.

difference electron density map

An **electron density map** constructed from **Fourier coefficients** based on the difference of observed and calculated **structure factor amplitudes**. Missing parts of the protein model show up as positive difference density, incorrectly placed parts as negative difference density.

difference Patterson map

A **Patterson map** constructed from **Fourier coefficients** based on the difference between **isomorphous difference** intensities or **anomalous difference** intensities. Reveals the interatomic distances between heavy or anomalous **marker atoms** and is used for marker atom **substructure** solution.

differentiation (M)

The process of finding a **derivative** of a function is called differentiation.

diffraction

Scattering discrete directions of electromagnetic radiation or particles by periodic assemblies.

diffraction-component precision index (12-101)

Measure for the coordinate precision based on *R*-free, data-to-parameter ratio, completeness, and resolution.

diffraction limit

The extent in solid angle to which diffracted X-rays can be observed. Generally expressed as a function of the (smallest) interplanar **lattice spacing**, and termed (highest) **resolution**, d_{min} . A direct measure for the information content of a diffraction pattern (Figures 1-6 and 9-8). Lattice spacing and **diffraction** angle are related through the **Bragg equation**. High resolution means small interplanar *d*-spacing.

diffraction image, pattern

A 2-dimensional image of diffracted X-rays or **X-ray reflections**, generally recorded on an area detector.

diffraction pattern symmetry – see Laue symmetry**diffraction spots – see X-ray reflections****diffractometer**

Instrument to generate X-rays and record diffracted X-rays. Minimally requires an X-ray source, X-ray optics, a goniostat, and an X-ray detector.

diffuse scattering – see thermal diffuse scattering**diffusion**

Net drift of molecules in the direction of lower concentration due

to random movement. A slow process in solution and even more so in solvent channels of crystals.

dihedral angle

The dihedral angle between two planes ABC and BCD of four consecutive atoms ABCD. **Torsion angles** are a special class of dihedral angles.

dimer

Molecular complex of two monomers (molecules). See also **homodimer**.

dimethyl sulfoxide, DMSO

A highly polar aprotic solvent, $(\text{CH}_3)_2\text{SO}$, that dissolves both polar and nonpolar compounds and is miscible in a wide range of organic solvents as well as water. Highly concentrated solutions of various small molecule ligand molecules in DMSO are used to introduce ligands with poor water solubility into crystallization drops for ligand **soaking**.

dipole interactions

Uncharged interactions between molecules, regions of molecules, or side chains that carry a dipole moment.

dipole moment

Represented by a dipole vector **P** and caused by charge separation or non-uniform charge distribution.

direct methods

Methods of crystallographic phase determination exploiting relations between strong normalized structure factors with certain index relationships. Routinely used in small molecule crystallography; in macromolecular crystallography predominantly for **substructure** solution.

direct rotation function

A rotation function based on the direct correlation between observed Patterson map and search model map. See also **Patterson correlation function**.

direct space – see real space

directed evolution

Method of random mutagenesis under selective pressure for specific desired properties such as solubility or crystallizability of the gene product.

direction cosine matrix

Matrix describing an arbitrary rotation through successive rotations about the three **Eulerian angles**. The determinant of a proper **direction cosine** matrix describing a rotation is 1 and the trace equals $1 + 2 \cos \kappa$, which is practical to compute the **principal Euler angle**.

direction cosines (A-26)

dispersion

The change of a physical property such as polarizability, refractive index, or atomic scattering factor with wavelength. Passing through an **X-ray absorption edge**, the **atomic scattering factor** changes rapidly with X-ray wavelength and becomes complex, leading to anomalous dispersion or **anomalous scattering**.

dispersive atom, marker

An atom that provides **anomalous scattering** contributions, used for anomalous phasing methods.

dispersive differences

Dispersive differences exist between the diffracted intensities of pairs of the same reflection recorded at two different wavelengths.

displacement parameter – see atomic displacement parameter

dissociation constant (K_d)

Measure of the tendency of a complex to dissociate. For components A and B and the binding equilibrium $A + B \leftrightarrow AB$, the dissociation constant is given by $[A][B]/[AB]$; the tighter the binding between A and B, the smaller the dissociation constant. The dissociation constant K_d is the reciprocal of K_a , the **affinity constant**.

disulfide bond (–S–S–)

Covalent linkage formed between two **thiol** groups on cysteines,

target bond length 2.03 Å. For extracellular proteins, a common way of joining two proteins together or linking different parts of the same protein. Formed in the endoplasmic reticulum of eukaryotic cells. Under reducing conditions during overexpression in bacteria, S–S bonds generally do not form, and special bacterial strains for the production of disulfide bond containing proteins have been developed.

DNA (deoxyribonucleic acid)

Polynucleotide formed from covalently linked deoxyribonucleotide units. The store of hereditary information within a cell and the carrier of this information from generation to generation. Common canonical form is B-DNA (Figure 2-42). A-DNA is a dehydrated form, and Z-DNA is a third form, not observed *in vivo*.

DNA ligase

Enzyme that joins the ends of two strands of DNA together with a covalent bond to make a continuous DNA strand.

DNA polymerase

Enzyme that synthesizes DNA by joining nucleotides together using a DNA template as a guide. Thermostable, high-fidelity proofreading polymerases are used for **polymerase chain reaction**.

DNA shuffling

A form of synthetic, *in vitro* gene recombination, often combined with **directed evolution**.

domain (C)

Region of crystal containing a certain well-defined orientation of molecules, separated from other domains by domain boundaries. See also **mosaicity**.

domain (protein domain)

Portion of a protein that has an independently folding **tertiary structure**, often with a specific function. Larger proteins are generally composed of several domains, each connected to the next by short flexible regions.

domain rotation function

Cross-rotation function that uses phase information from NCS related electron density domains to determine NCS rotations.

domain swapping

Occurs in dimeric proteins when one part or domain of one molecule exchanges position with the same part or domain of the other molecule.

domain truncation

Separating a multi-domain protein into smaller, individually expressed domains and modifying the domain boundaries by recombinant DNA techniques or biochemical proteolysis.

dose mode

Data collection mode where the total incident radiation over each exposure increment (frame) is constant.

dot product – see scalar product

double helix

The 3-dimensional structure of DNA, in which two antiparallel DNA chains, held together by hydrogen bonding between the bases, are wound into a helix. Figure 2-44.

dual space (direct) methods

Substructure solution methods, Patterson-seeded in the case of **SHELXD**, that cycle between reciprocal space tangent refinement and direct space peak picking.

dual space methods

General term for iterative crystallographic methods that cycle between reciprocal space and real space; such as in direct methods or dummy atom placement.

dummy atom placement (refinement)

Dual space method for phase improvement and model building based on iterative placement, refinement, and removal of dummy atoms. Implemented in **ARP/wARP**.

duplex DNA

Double-stranded DNA.

dyad (C)

Name for 2-fold **rotation axis**.

dyad product (M, A-43)

Square matrix obtained by multiplication of column vector with row vector.

dynamic light scattering, PCS, QELS

Single angle light scattering method for particle size determination that measures the scattered photon correlation. Also called photon correlation spectroscopy (PCS) or quasi-elastic light scattering (QELS).

E

E. coli – see *Escherichia coli*

E-squared minus one, $\langle |E^2 - 1| \rangle$ – see **mean *E*-squared minus one**.

E-values – see **normalized structure factors**

edge scan – see **X-ray excitation scan**

effective resolution

Resolution d_{min} of diffraction data, corrected for completeness C :
 $d_{eff} = d_{min} \cdot C^{-1/3}$.

eigenvalue

Scalar value λ associated with a non-zero **eigenvector** of a linear matrix equation in the form of an **eigenvalue equation**.

eigenvalue equation

Linear system of equations $\mathbf{Ax} = \lambda\mathbf{x}$ where \mathbf{A} is a matrix describing a linear transformation (such as the **direction cosine matrix** of a **rotation**) with \mathbf{x} a nonzero **eigenvector** and λ the scalar **eigenvalue**.

eigenvector (eigenfunction)

Solution vector (function) of linear matrix equation, associated with a given **eigenvalue**.

elastic scattering

Scattering process in which no energy transfer occurs and **wave-length** or energy is maintained. The discrete Bragg scattering exploited in X-ray diffraction is elastic.

electromagnetic radiation

Described in dual picture as either (i) transversal electromagnetic wave with electromagnetic field vector and magnetic field vector perpendicular to each other and to the propagation direction; or (ii) an uncharged **photon** particle with no rest mass of energy $E = h\nu$ with h the Planck constant and ν the frequency. See Box 6-1 and **energy-wavelength conversion**.

electromagnetic wave – see **electromagnetic radiation**

electron

Subatomic particle that carries a negative electric charge. Electrons, together with atomic nuclei made of protons and neutrons, make up the atoms. Electrons scatter X-rays, and the **Fourier transform** of the **diffraction** patterns in **reciprocal space** therefore yields the **electron density** of the scattering object in **real space**.

electron density

The **Fourier transform** of experimentally determined **structure factor amplitudes** (proportional to the square root of diffraction spot intensities) plus additional **phases** (from phasing experiments) as **Fourier coefficients** generates the electron density of the diffracting objects (molecules).

electron density map

A representation of the **electron density** in the form of a 2-dimensional contour plot or a 3-dimensional grid connecting points or voxels of equal electron density. The level of density is often expressed in terms of **standard deviations** (σ) above the mean density. A high density level is represented for example by 5σ contours, while a low density level is at a 1σ level.

electrophoresis

Technique for separating molecules (typically proteins or nucleic

acids) on the basis of their speed of migration through a porous medium when subjected to a strong electric field.

electroporation

Method for introducing DNA into cells, especially bacteria, in which a brief electric shock makes the cell membrane temporarily permeable to the foreign DNA.

empirical potentials

Expression of deviation from stereochemical target values as an energy potential function.

enantiomorph

One of the two partners in an enantiomorph pair having opposite **chirality** or handedness.

enantiomorph ambiguity

Both **enantiomorphs** of a substructure produce the same **Pattern map**.

enantiomorph pairs of space group

Twenty-two of the 65 **chiral space groups** with **screw axes** $3_1, 3_2; 4_1, 4_2; 6_1, 6_2; 6_3, 6_4$ form 11 enantiomorph pairs. The pairs are $P3_1, -P3_2$, $P3_1, 21-P3_2, 21$, $P3_1, 12-P3_2, 12$, $P4_1, -P4_3$, $P4_1, 22-P4_3, 22$, $P4_1, 2, 2-P4_3, 2, 2$, $P6_1, -P6_5$, $P6_2, -P6_4$, $P6_1, 22-P6_5, 22$, $P6_2, 22-P6_4, 22$, and cubic $P4_1, 32-P4_3, 32$. When the handedness of the (sub)structure is changed, the screw axis must also be changed to the other **enantiomorph**. See also **substructure handedness**.

endoglycosidase

Enzyme that cleaves **oligosaccharides** *within* the polysaccharide chain; used for enzymatic deglycosylation of **glycoproteins**.

endonuclease

Enzyme that cleaves nucleic acids *within* the polynucleotide chain. *Compare* **exonuclease**.

endoplasmic reticulum (ER)

Labyrinthine membrane-bounded compartment in the cytoplasm of eukaryotic cells, where lipids' membrane-bound proteins, and secretory proteins are synthesized.

energy refinement

Refinement in which the X-ray and restraint SRS residuals are expressed as potential energy terms. Minimized commonly by simulated annealing.

energy-wavelength conversion

Energy in electronvolts (eV) = $12397.639/\text{wavelength in } \text{\AA}$.

enthalpy, *H*

Change in enthalpy is the heat transfer ΔH during a chemical or physical process. *Compare* **free energy**.

entropy, *S*

The change in entropy ΔS is a measure of the change in order or degrees of freedom during a chemical or physical process. ΔS is negative for a process in which order is created, as in formation of an ordered crystal. The associated energy term is $-T\Delta S$. *Compare* **free energy**.

enzyme

Protein that catalyzes a specific chemical reaction.

enzyme-coupled receptor

A major type of cell-surface receptor (membrane protein) that has a cytoplasmic domain that either has enzymatic activity or is associated with an intracellular enzyme. In either case, the enzymatic activity is stimulated by an extracellular ligand binding to the receptor. Typical example GPCRs, G-protein coupled receptors.

epitaxy

Growth of crystals on a crystalline substrate acting as nucleation site with a matching or specific relation of cell dimensions between growing crystal and crystalline substrate.

epitope

Specific region (antigenic determinant) of an antigen that binds to an **antibody** or a T-cell receptor.

epitope-binding regions

Hypervariable complementarity determining regions (CDRs) of

an antibody binding to antigen.

epsilon-factor (epsilon-zone)

A correction factor applied to **intensity** statistics accounting for overlap in reciprocal space of certain symmetry related reflections.

equipoints – see **equivalent positions**

equivalent positions

Positions in a crystal structure that are symmetry related and generated by the space group symmetry operations.

equivalent reflections

X-ray diffraction spots or structure factor amplitudes related by reciprocal space symmetry.

error function, erf(x)

Sigmoid function obtained by integration from 0 to x of squared negative exponential functions, such as the normal distribution.

Escherichia coli (E. coli, E.C.)

Rod-like bacterium normally found in the colon of humans and other mammals. Engineered strains are widely used as heterologous overexpression hosts.

eucentric point

Unique point in a diffractometer setup where all goniostat axes as well as the X-ray beam and optical centering microscope axis intersect.

eukaryote (eucaryote)

Organism composed of one or more cells that have a distinct nucleus. Member of one of the three main divisions of the living world, the other two being bacteria and archaea. *Compare prokaryote.*

Euler angle – see **principal Euler angle**

Euler axis

The Euler axis is the single, unique axis around which any general **rotation** is defined. It is obtained as the **eigenvector** belonging to the real **eigenvalue** of 1 when solving the linear matrix equation represented by the **direction cosine matrix R** describing the rotation. The **principal Euler angle** defines the single rotation angle around the Euler axis, and is equivalent to the rotation angle κ in **spherical coordinate** representation.

Eulerian angles

The three angles used in the **direction cosine matrix** to describe a rotation. Multiple conventions exist to define these three angles (Sidebar 11-1). Different from **principal Euler angle**.

Euler's formula (6-5; Section A.5)

evolutionary programming

Genetic algorithm that does not apply crossover and combination of parent properties but works well in continuous cases.

Ewald sphere

A sphere with radius of $1/\lambda$, with the diffracting crystal in its center, used to visualize diffraction geometry. Whenever a **reciprocal lattice point** intersects the Ewald sphere, the diffraction conditions in the form of the **Laue** and **Bragg equations** are fulfilled.

EXAFS – extended X-ray absorption fine structure

Region of the X-ray absorption spectrum above the absorption edge. Typical features are EXAFS wiggles representing a Fourier transform of nearest neighbor geometry and distances.

excluded volume effect

General effect in macromolecular solutions increasing the effective concentration of macromolecules, because one part of a macromolecular chain cannot occupy space that is already occupied by another part of the same molecule.

exon

Segment of a eukaryotic gene that consists of a sequence of nucleotides that will be represented in mRNA or other mature RNA molecule. In protein-coding genes, exons encode the amino acids in the protein. An exon is usually adjacent to a noncoding DNA segment called an **intron**.

exonuclease

Enzyme that cleaves nucleotides one at a time from the ends of polynucleotides. *Compare endonuclease.*

expectation value (7-42)

In general terms, the integral of a random variable with respect to its probability measure. For discrete random variables this is equivalent to the weighted sum of the possible (or observed) discrete values forming the discrete probability distribution. For continuous random variables with a **probability density function** it is the probability density-weighted integral of the function values. Annotated also as $E(x)$ or $\langle x \rangle$.

experimental design

Formal statistical design of an experiment or study so that unbiased and valid statistical analysis of the outcome is possible.

experimental phasing

Encompasses all crystallographic **phasing methods** that provide phase information independent of any starting model. *Compare molecular replacement.*

exponential bulk solvent model

Bulk solvent correction valid for very low resolution reflections only, based on **Babinet's principle**.

expression (B)

Process by which information from a **gene** is used in the synthesis of a functional gene product, generally a **protein**. Used also for **overexpression** of proteins in various **expression hosts**.

expression host

Organism in which **heterologous overexpression** of a protein takes place. Most commonly bacteria, yeasts, insect cells, and cell lines of higher organisms.

expression vector

A **virus** or **plasmid** that carries a DNA sequence into a suitable **expression host** cell and there directs the synthesis of the protein encoded by the DNA sequence.

extensive parameters

In thermodynamics, extensive parameters are those explicitly depending on the amount of material. *Compare intensive parameters.*

extinctions (C) – see **systematic absences**

extrema (M)

Maxima or minima of a function or probability distribution.

F

face centered

Lattice or unit cell centering with (Bravais) centering vector $F = (\frac{1}{2} \frac{1}{2} 0)$, $(\frac{1}{2} 0 \frac{1}{2})$, and $(0 \frac{1}{2} \frac{1}{2})$. Possible in orthorhombic (*oF*) and cubic (*cF*) lattices.

factorial design

A type of **experimental design** balancing the occurrence of possible factors (reagents, pH, drop size/ratio), their factor levels, and their combinations during the sampling process.

fast Fourier transform (FFT)

Method to rapidly compute Fourier transforms on $n = 2^k$ grid points.

fast rotation function (11-32)

A function calculating the overlap of Patterson functions in reciprocal space rapidly by expressing the **reciprocal space interference function** in spherical functions, with specially parameterized Euler angles. Developed originally by Crowther and Blow, and implemented in modern variants in various molecular replacement packages.

fast translation function (11-38)

A Patterson **translation function** that computes fast due to subtracting intermolecular vectors from the **Patterson map**, developed originally by Crowther and Blow, and implemented in modern variants in various molecular replacement packages.

fatty acid

Carboxylic acid with a long hydrocarbon tail.

fermentation

Anaerobic energy-yielding metabolic pathway.

figure of merit, fom, m

A statistic for the probability of a **phase angle** to be correct, expressed as mean phase error $m = \langle \cos(\Delta\varphi) \rangle$.

fine slicing

Method of recording diffraction images in small rotation increments, generally in the range of 0.1 to 0.3°. Allows 3-dimensional diffraction spot profile fitting.

fission yeast

Common name for the yeast model organism *Schizosaccharomyces pombe*. It divides to give two equal-sized cells.

Flack parameter

The Flack parameter x is defined in the range $0 \leq x \leq 1$ as $F_{h,x}^2 = (1-x)F_h^2 + xF_{-h}^2$. If the handedness is chosen correctly, x is close to 0; if the structure needs to be inverted, x is close to 1. This can give useful indications during heavy atom **substructure** refinement about substructure **handedness**. During refinement of protein structures strange Flack parameters may indicate **microscopic twinning**.

flash cooling

Rapid cooling of protein crystals by direct transfer into a **cryogenic** environment (generally **liquid nitrogen**) that prevents ice formation in the **mother liquor** surrounding a **harvested** protein crystal. The mother liquor solidifies in **vitreous** or **amorphous** form during flash cooling. Generally requires some form of **cryoprotection**.

flat bulk solvent model

Bulk solvent correction that generates a solvent mask around the protein and then fills the solvent regions with continuous density.

fluorescence

Emission of radiation by atoms when excited electrons return to their energy ground state, generally within nanoseconds. *Compare phosphorescence*.

fluorescence scan – see X-ray excitation scan**flux (C) – see photon flux****fold**

A specific 3-dimensional arrangement of **secondary structure** elements of a protein forming a specific and stable **tertiary structure** that belongs to a certain fold family of structurally similar proteins.

form factor (C) – see atomic scattering factor**Fourier coefficients**

The coefficients in a **Fourier integral** or series. In crystallography, the **complex structure factor**; or the **structure factor amplitude** and its associated **phase angle**.

Fourier convolution theorem (9-15)

States that the Fourier transform (FT) of a **convolution** of two functions equals the product of the individual FTs.

Fourier indexing

Indexing method that exploits **Fourier transforms** of the diffraction spots in certain directions to assign lattice vectors.

Fourier integral, series (9-2)

A complex integral or infinite complex series expressing a function as sum of cosine and sine terms or their equivalent complex exponential representation.

Fourier synthesis, summation (9-13, 9-14)

Procedure of carrying out the Fourier integration or Fourier summation. In crystallography, the result is either complex structure factors in reciprocal space or electron density in real space, which are Fourier transforms of each other.

Fourier transform (9-5)

A bijective transformation of a function from one domain into its reciprocal domain by means of Fourier integration or summation. In crystallography, the transformation of **complex structure factors** in **reciprocal space** into **electron density** in **real space** and *vice versa*.

Fourier transform interpolation

Fastest method to calculate structure factors, primarily used, for example, in **stochastic search** algorithms.

Fourier truncation ripples

Result around (heavy) atoms when truncated (phased) data are used for Fourier reconstruction of electron density.

fractional coordinates

Position of a point in a crystal structure expressed in dimensionless fractions x, y, z of the **unit cell vectors a, b, c**.

fragment screening – see crystallographic fragment screening**frame – see diffraction image****free energy, G**

Also Gibbs energy, the driving force for chemical reactions. The change in free energy ΔG must be negative for a reaction to spontaneously occur. At the equilibrium, the free energy is zero. It is the sum of an enthalpic term and an entropic term, $\Delta G = \Delta H - T \Delta S$.

free-interface diffusion

Crystallization technique where the crystallization cocktail and the protein solution freely diffuse against each other. Carried out in capillaries or in **microfluidic** chips.

free likelihood ratio

Measure for statistical significance of differences between models. *See also Bayes factor*.

Free Lunch

A method of **phase extension** initially replacing unobserved high resolution reflections with those calculated by George Sheldrick's **sphere of influence algorithm** (Sidebar 10-10).

free R-value – see cross-validation R-value (12-35)**free radical**

Atom or molecule which is extremely reactive due to at least one unpaired electron. Responsible for intracellular DNA damage *in vivo* and in part for the radiation damage to protein crystals during exposure to X-rays.

frequency

Common symbol ν , related to energy by $E = h\nu$ with h the Planck constant. *See also wavelength–energy conversion*.

Friedel pair

The reflection **h** and its Friedel opposite or mate **-h** form a Friedel pair. Friedel pairs are also **Bijvoet pairs**, but not all Bijvoet pairs are Friedel pairs. For general acentric reflections, **anomalous differences** exist between members of a Bijvoet pair.

Friedel's law

States that in the absence of **anomalous scattering** contributions, $F_h = F_{-h}$. The **structure factor amplitudes** (or measured intensities) of reflections with conjugate phase (i.e. the members of a Friedel pair) are identical. In the presence of anomalous signal, Friedel's law breaks down, and the relation $F_h = F_{-h}$ remains true only for **centric reflections**.

full matrix refinement (minimization)

Minimization algorithm in which all first and second derivative matrix elements are evaluated.

full reflection

An X-ray reflection or reciprocal lattice point that appears entirely on a single diffraction image or frame.

full width at half maximum (FWHM)

The width across a peak at half of its maximum. Particularly useful for cases such as Lorentz or Cauchy (spectral) lines where the variance is undefined. Figure 7-3.

function (M)

Arithmetic expression describing the dependence between two or more quantities, one or more of which are known (independent variables) and the others which are generated. Example: $y = f(x) = x^2$ means that y , a dependent variable, which is a function of the independent variable x , is given by the square of x .

fungi

Eukaryotic organisms that include the yeasts, molds, and mushrooms. Many plant diseases and a relatively small number of animal diseases are caused by fungi.

fusion protein, peptide

Engineered protein that combines two or more normally separate polypeptides, expressed from recombinant gene. Most commonly **affinity tags** or solubility enhancers.

G**G-function** – see **reciprocal space interference function****G-protein-coupled receptor (GPCR)**

A cell-surface receptor with seven transmembrane helices that, when activated by its extracellular ligand, activates a G-protein, which in turn, activates either an enzyme or ion channel in the plasma membrane.

ganglioside

Any glycolipid having one or more sialic acid residues in its structure. Found in the plasma membrane of eukaryotic cells and especially abundant in nerve cells.

Gaussian normal distribution (7-24)

A ubiquitous **probability distribution function**, due to the **central limit theorem**. The normal distribution is symmetric about its maximum representing the mean or **expectation value**, with its **variance** defined as a squared residual. Figure 7-3 and 7-6.

gel-shift assay (gel-mobility shift assay)

Technique for (i) detecting proteins bound to a specific DNA sequence by the fact that the bound protein slows down the migration of the DNA fragment through a gel during gel electrophoresis, or (ii) detecting ligands or heavy atoms bound to proteins by slowing down protein migration in a native, non-denaturing SDS-PAGE gel.

gene

Region of DNA that is transcribed as a single unit and carries information for a discrete hereditary characteristic, usually corresponding to (i) a single **protein** (or set of related proteins generated by variant post-transcriptional processing), or (ii) a single RNA (or set of closely related RNAs).

general position (C)

A point position in a crystal structure with **fractional coordinates** x, y, z that is not located on any **symmetry element**.

generator (C)

A symmetry operation that is one of the one to three symmetry operations combined with the lattice translations into a **space group**.

genetic algorithms

Stochastic search or optimization methods that generate and recombine solutions from random starting values following the principles of genetic evolution. *Compare evolutionary algorithms.*

genetic engineering – see **recombinant DNA methods****geometry term**

Sum of residuals squared term (SRS) for the deviations of model geometry from independently determined geometry target values.

GIGO principle

Universal scientific principle stating that in any given system, garbage in equals garbage out.

glide plane

A combination of **mirror operation** with translation. Not allowed for **chiral** motifs.

glutamate, glutamic acid

Charged, acidic L- α -amino acid with $-\text{CH}_2\text{CH}_2\text{COO}^-$ side chain. Hydrogen bond acceptor.

glutamine

Polar L- α -amino acid with $-\text{CH}_2\text{CH}_2\text{CONH}_2$ side chain. Hydrogen bond donor.

glutaraldehyde

A dialdehyde $\text{HOC}-(\text{CH}_2)_3-\text{COH}$ used as a cross-linking reagent for protein crystals.

glycine

Smallest, non-chiral α -amino acid with a second H atom instead of a side chain branch.

glycoprotein

Any **glycosylated** protein with one or more oligosaccharide chains covalently linked to amino-acid side chains. Most secreted proteins and most proteins exposed on the outer surface of the plasma membrane are glycoproteins. *See also glycosylation.*

glycosylation

Covalent attachment of **carbohydrate** moieties, most frequently to Asn (N-glycosylation) or Ser and Thr (O-glycosylation). Can cause problems in crystallization due to conformational and chemical inhomogeneity. A **posttranslational modification** that does not occur in bacterial **expression hosts**.

goniometer – see **goniostat****goniometer head**

Small precision instrument allowing translational adjustment of the crystal for centering, mounted with standardized thread on the **goniostat**. Some eucentric goniometer heads also have adjustable arcs.

goniostat

Precision device allowing orienting a crystal so that it remains centered in the **eucentric point** of a diffractometer. Goniostats can be simple single-axis instruments or complex multi-axis (multi-circle) instruments.

gradient (descent) minimization (optimization)

Multivariate optimization algorithms that use in various implementations the first derivative to determine parameter shifts.

green fluorescent protein

A β -barrel structure harboring a chromophore, used in many variants as a fluorescence label.

grid screen

An exhaustive screening protocol varying two experimental parameters in a 2-dimensional matrix (grid) design.

group (M)

An abstract algebraic structure, consisting of a set (of objects or numbers, for example) and some operators acting on the members of the set. The conditions for set and operators forming a group are (i) **closure**—any operation can only generate a member of the group, (ii) **identity**—one and only one **identity operation** exists, (iii) **inversion**—in the mathematical sense that reversal of the operation generates the original object, and (iv) **associativity**—that $a(bc) = (ab)c$ (which does not automatically apply **commutativity**, that is, $ab \neq ba$). **Cyclic groups** are commutative.

guanidinium, guanidyl

Chemical group $-\text{NHC}(\text{NH}_2)_2^+$, derived from guanidine, in side chain of arginine.

H**habit (C)** – see **crystal habit****handedness (M)**

Refers to the relative orientation of **basis** vectors. In the 3-dimensional case, the basis $[0, \mathbf{a}, \mathbf{b}, \mathbf{c}]$ is right-handed if the vectors $\mathbf{a}, \mathbf{b}, \mathbf{c}$ follow the sequence of thumb, index finger, middle finger of the right hand. Figure 5-25.

handedness (C)

Refers to one of the two configurations of a **chiral center**, chiral molecule, or **asymmetric** atomic assembly or marker atom **substructure**. *See also* **chirality**.

handedness ambiguity – *see* **substructure handedness**

hanging drop (C)

Crystallization method by **vapor diffusion** where a drop of protein solution plus precipitate is placed over a well solution of **precipitate** solution. *See also* **sitting drop**.

Harker diagram

Graphic representation of the **phasing equations**, where the **structure factor amplitudes** are represented as circles.

Harker sections

Certain sections (space-group-dependent **centric projections**) of a **Patterson map**, which reveal interatomic distance vectors between symmetry related atoms. Used in marker atom **substructure** solutions for **experimental phasing**.

harvesting (C)

The process of removing a crystal from the growth solution or **mother liquor**, most frequently with some form of harvesting loop.

heat shock protein (Hsp, stress-response protein)

Large family of highly conserved molecular chaperone proteins, so named because they are synthesized in increased amounts in response to an elevated temperature or other stressful treatment. Hsps have important roles in aiding correct protein folding or refolding.

Heaviside function

A unit step function, with Fourier transform of $\sin(x)/x$.

heavy atom

An atom of a heavy element (generally a heavy metal ion soaked into a protein crystal) used as marker atom in **experimental phasing**. Most heavy atoms also are **dispersive markers** providing **anomalous signal** for phasing.

heavy atom derivative (crystal)

Obtained by soaking **heavy atoms** into protein crystals or by co-crystallizing a protein with heavy atom compounds. In order to be useful for **experimental phasing**, the heavy atom derivative crystal needs to be **isomorphous** to the native crystal.

heavy atom refinement

Minimization of **sum of residuals squared** ($F_{\text{PH}}(\text{obs}) - F_{\text{PH}}(\text{calc})$)² by adjusting the parameters of the **heavy atom substructure**, generally by **maximum likelihood** methods, in order to obtain improved protein phases. Can be omitted if the substructure solution is (i) correct and (ii) complete, frequently the case for **dual space** substructure solution **methods** such as **SHELXD** or **Shake and Bake**.

heavy chain (H chain)

The larger of the two types of polypeptide chain in an immunoglobulin molecule, extending into the F_c fragment.

helix-loop-helix

DNA-binding structural motif present in many gene regulatory proteins, consisting of a short alpha helix connected by a flexible loop to a second, longer alpha helix. Distinct from the helix-turn-helix motif.

helix-turn-helix

DNA-binding structural motif present in many gene regulatory proteins, consisting of two alpha helices held at a fixed angle and connected by a short chain of amino acids, constituting the turn. Proteins containing this motif frequently form symmetric dimers and bind to DNA sequences that are themselves similar and arranged symmetrically.

hemihedral twinning

Mero-hedral twin with a single **twin operator**. If the twin fraction of a single twin operator is 0.5, it forms a perfect hemihedral twin with exactly equal amounts of each domain orientation (Figure 8-18). A corresponding projection of the diffraction pattern acquires additional perfect symmetry.

Hendrickson-Lattman coefficients (10-82)

Used for phase combination where each phase probability is expressed in terms of $\cos \phi$, $\sin \phi$, $\cos 2\phi$, and $\sin 2\phi$ and coefficients A , B , C , D (plus a normalization constant K) whose form depends on the particular **phase probability distribution**.

Hermann-Mauguin symbol – *see* **space group symbol**

Hermitian matrix (M)

A Hermitian matrix (or self-adjoint matrix) is a **square matrix** with complex elements which is equal to its own conjugate **transpose**. **Eigenvalues** of a Hermitian matrix are real with **orthogonal eigenvectors**. Physical observables obtained via Hermitian operators are real.

Hessian matrix

The Hessian (matrix) is the square matrix of second-order partial derivatives of a function; that is, it describes the local curvature of a multivariate function.

heterodimer

Molecular complex, **heterooligomer**, composed of two different **monomers**; protein complex composed of two different **polypeptide** chains. *Compare* **homodimer**.

heterogeneous nucleation – *see* **nucleation**

heterologous overexpression

Expression of high levels of a protein in a host system different from the native organism.

heterooligomer

Molecular complex composed of multiple different **monomers**; protein complex composed of multiple different **polypeptide** chains. *Compare* **homooligomer**.

heteropolymer

Macromolecule assembled from a subset of different building blocks or **monomers**. Proteins are heteropolymers made of **proteinogenic amino acids**. *Compare* **homopolymer**.

hexagonal

Lattice or crystal structure with cell parameters $a = b$, c , $\alpha = \beta = 90^\circ$, $\gamma = 120^\circ$, with primitive (hP) lattice type and 6-fold internal minimal symmetry.

high energy remote data set

Anomalous diffraction data collected about 100 eV or so above the absorption edge. Intended to maximize dispersive signal to **inflection data set**.

high resolution

Higher resolution means higher diffraction angle, equivalent to smaller sampled interplanar d -spacing d_{min} , implying finer detail discernible in the molecular structure. Figure 1-6.

histidine

Aromatic amino acid with imidazole ring in side chain $-\text{CH}_2(\text{C}_3\text{N}_2\text{H}_3)^+$, weakly protonated at physiological pH.

histogram matching (C)

A **solvent modification** method providing phase improvement and **phase extension** based on matching the initial electron density distribution to that observed for actual protein structures.

homeodomain

DNA-binding domain that defines a class of gene regulatory proteins important in development of higher organisms.

homodimer

Molecular complex, **homooligomer**, composed of two identical **monomers**; protein complex composed of two identical **polypeptide** chains. *Compare* **heterodimer**.

homogeneous nucleation – *see* **nucleation**

homolog

One of two or more genes that are similar in sequence as a result of derivation from the same ancestral gene. The term covers both **orthologs** and **paralogs**. Homology often but not exclusively indicates structural similarity.

homology modeling

Method of comparative modeling based on experimental structure template of a related structurally similar model as identified by sequence alignment.

homooligomer

Molecular complex composed of multiple identical **monomers**; protein complex composed of multiple identical **polypeptide** chains. *Compare* **heterooligomer**.

homopolymer

Macromolecule assembled from a subset of identical building blocks. *Compare* **heteropolymer**.

homozygous

Individual who carries two identical alleles of a gene affecting a given trait on the two corresponding homologous chromosomes.

HSQC

Heteronuclear single-quantum coherence, 2-dimensional NMR technique, separating backbone amides according to their ^1H and ^{15}N resonance frequencies.

hybridoma

Cell line used in the production of monoclonal **antibodies**, obtained by fusing antibody-secreting B cells with cells of a B-lymphocyte tumor.

hydrogen bond

The most important bond type in biological systems and processes. Ubiquitous noncovalent, semi-directional bond in which an electropositive hydrogen atom is partially shared by two electro-negative atoms, one acting as a hydrogen bond donor and the other as hydrogen bond acceptor. *See also* **bifurcated hydrogen bond**.

hydrodynamic radius

Radius of a macromolecule in solution determined by the Stokes–Einstein relation (4-4).

hydronium ion

Solvated **proton**, H_3O^+ . The form generally taken by protons in aqueous solution.

hydrophilic

Dissolving readily in water. Literally, “water loving.”

hydrophobic (lipophilic)

Not dissolving readily in water. Literally, “water hating.”

hydrophobic force

Force exerted by the hydrogen-bonded network of water molecules that brings two nonpolar surfaces together by excluding water between them.

hydrophobic moment

The hydrophobic moment reflects the periodicity of hydrophobicity of a peptide, as measured per residue for a specified angle of rotation.

hydroxyl

Chemical group consisting of a hydrogen atom linked to an oxygen (–OH), as in an alcohol.

hygroscopic

A hygroscopic substance attracts water molecules from the surrounding environment.

I**I-centered** – *see* **body centered****identity matrix**

An identity matrix has unity values in the diagonal elements and zeros elsewhere.

identity operation

An operation that maps an object onto itself; described by an **identity matrix**.

imaging plate

Area detector that stores X-ray photon energy in a **phosphorescent** material, read out later optically by laser excitation.

improper non-crystallographic symmetry (NCS)

General non-crystallographic symmetry that is both rotational and translational, does not comply with closed **group** limitations. *Compare* **proper non-crystallographic symmetry**.

inclusion body

Inclusion bodies are dense proteinaceous aggregates that form in the cytoplasm, notably when the protein folding machinery becomes overwhelmed. Inclusion bodies generally contain misfolded proteins.

incommensurately modulated structure

Crystal structure in which motifs vary in a systematic fashion out-of-phase or incommensurate with the periodic translation period of the original lattice. Manifests itself in reciprocal space through **satellite reflections** next to the Bragg peaks of the parent cell. *Compare* **commensurately modulated structure**.

indexing

The consistent assignment of three linearly independent basis vectors to a reciprocal or real lattice, generally by **Fourier indexing** methods.

indexing possibilities

The combination of the 14 **Bravais lattice** types and axis permutation generates 44 possible indexing choices. The highest symmetry indexing with the smallest penalty score is generally the correct choice.

inducible promoter

A regulatory DNA sequence that allows expression of an associated gene to be switched on by a particular molecular or physical stimulus.

inductive inference (logic)

An argument is inductive when the truth of the conclusion is given with a certain probability based on experimental or prior knowledge, that is, its corresponding conditional is a likelihood (function). *Compare* **deductive inference**.

inelastic scattering

Scattering process where the scattered radiation or particle loses or gains energy during the scattering process.

inflating the variance

Method to increase the variance of a **likelihood function** in order to implicitly account for additional (model) errors.

inflection data set

Anomalous diffraction data collected at the absorption edge inflection point, corresponding to the negative peak in the **Kramers–Kronig transform** of the **absorption edge** scan.

insertion device

Magnetic instrument that is inserted into the straight sections of an **electron synchrotron storage ring**, used to generate intense **synchrotron X-ray radiation**. **Wigglers** and **undulators** are insertion devices.

inside–outside distribution

Distribution of propensity of residues to be located either at the surface of protein or in the hydrophobic core of a protein. Useful in low-resolution X-ray structure validation and validation of predicted structures.

insulin

Polypeptide hormone that is secreted by B-cells in the pancreas regulating glucose metabolism in animals.

integral extinctions – *see* **systematic absences****integral membrane protein**

Membrane protein that is embedded in the lipid bilayer with transmembrane helices so that it does not have independently stable or expressible ectodomains or cytosolic domains.

integrating (C)

Reading of detector pixels and combining them to raw reflection intensities.

intensity, of diffracted X-rays (6-49, 6-52)

The intensity of an X-ray reflection or diffraction spot is proportional to the square of the corresponding structure factor amplitude.

intensity distribution

Probability distribution functions for diffraction intensities, commonly expressed in normalized form as squared normalized structure factors (E^2 -values). *See also* **normalized structure factors** (E).

intensive parameters

In thermodynamics, intensive parameters are those not depending on the amount of material in the system, such as pH or temperature. *Compare* **extensive parameters**.

interatomic distance vectors

Distances between atoms in a structure (cross-vectors) and symmetry related atoms (self-vectors) in a crystallographic unit cell, comprising the Patterson space. *See also* **Patterson function**.

interference function – *see* **reciprocal space interference function****interferon (IFN)**

Member of a class of cytokines secreted by virus-infected cells and certain types of activated T cells. Interferons induce antiviral responses, inhibiting viral replication and stimulating macrophages and natural killer cells to kill virus-infected cells.

internal symmetry

The symmetry of a **lattice** as defined by the internal arrangement of motifs, defining the **crystal system**.

interplanar distance vector – *see* **lattice spacing****interplanar spacing** – *see* **lattice spacing****intron**

Noncoding region of a eukaryotic gene that is transcribed into an RNA molecule but is then excised by RNA splicing during production of the mRNA or other functional RNA.

inverse probability – *see* **likelihood****inversion** (C)

Symmetry operation that inverts the position vector, that is, x becomes $-x$. Inversion changes the handedness of a **substructure** or the **chirality** of a molecule.

inversion (M)

Operation that generates a matrix M^{-1} so that $MM^{-1} = I$, the **identity matrix**.

inversion axis

A crystallographic **symmetry operation** combining **rotation** with **inversion**, not allowed on **chiral** motifs such as protein molecules.

inversion center

The unique point through which a structure is inverted. Generally an origin of a unit cell, but three exceptions to the inversion about the origin exist: In chiral space groups $I4_1$, $I4_122$, and $F4_132$, the origin is not located on the enantiomorph axis, and the center of inversion does not coincide with the origin. Their inversion operators are $(-x, \frac{1}{2}-y, -z)$, $(-x, \frac{1}{2}-y, \frac{1}{4}-z)$, $(\frac{1}{4}-x, \frac{1}{4}-y, \frac{1}{4}-z)$, respectively. *See also* **substructure handedness**.

ion

An atom that has either gained or lost electrons to acquire a negative (anion) or positive (cation) charge.

ion channel

Transmembrane protein complex that forms a water-filled channel across the lipid bilayer through which specific inorganic ions can diffuse down their electrochemical gradients.

ionizing radiation

High energy electromagnetic or corpuscular radiation causing bond breakage and formation of **free radicals**.

isoelectric focusing

A technique for separating different proteins by their electric

charge, on 2-dimensional gels for analytical purposes; also possible on preparative scale.

isoelectric point, pI

The pH value at which the net charge (the sum of all local charges) of a protein is zero. Also the point where protein **solubility** is minimal, and the protein does not migrate in an electric field.

isomer

Molecule formed from the same atoms and connectivities as another but having a different 3-dimensional conformation.

isoleucine

A **hydrophobic** L- α -amino acid with an asymmetrically branched aliphatic side chain $-\text{CH}(\text{CH}_3)\text{CH}_2\text{CH}_3$ and a second chiral center at C β , with absolute configuration (2S, 3R).

isomerase

Enzyme that catalyzes the rearrangement of bonds within a single molecule.

isomorphous (C)

Of the same crystal structure. Native structure and **heavy atom derivative** structure need to be isomorphous to generate **isomorphous difference** intensities useful for phasing. **Anomalous** and **dispersive difference** intensities are inherently isomorphous, subject to **radiation damage**. *See also* **RIP**.

isomorphous difference

Isomorphous differences exist between the diffracted intensities of a native protein and a **heavy atom derivative**.

isomorphous difference Patterson map

Patterson map formed from isomorphous difference coefficients $(F_{\text{PA}} - F_{\text{P}})^2$.

isomorphous replacement

A method of experimental phase determination based on determination of a heavy atom marker **substructure**. Requires both a native crystal and one or more **isomorphous derivative crystal(s)**. Based on the fact that the complex structure factors (but not the **amplitudes**) are additive: $F_{\text{PH}} = F_{\text{P}} + F_{\text{H}}$, or split into amplitude and phase term, $F_{\text{PH}} \cdot \exp(i\phi_{\text{PH}}) = F_{\text{P}} \cdot \exp(i\phi_{\text{P}}) + F_{\text{H}} \cdot \exp(i\phi_{\text{H}})$. F_{P} and F_{PH} as well as F_{H} (both F_{H} and ϕ_{H}) are known from the substructure solution. P; protein; PH, heavy atom derivative; H, heavy atom.

isotope

One of a number of forms of an element differing in atomic weight, that is, having the same number of protons and electrons but different neutron number. Isotopes are chemically equivalent with the exception of kinetic isotope effects caused by different atomic mass.

isotropic

Directionally independent, uniform.

isotropic displacement parameter – *see* **atomic displacement parameter****isotropic overall B-factor** – *see* **overall B-factor****J****joint probability**

The product of two or more, generally conditional, probabilities.

joule

Standard unit of energy in the meter-kilogram-second system. One joule is the energy delivered in 1 second by a 1-watt power source. Equal to 0.23901 calories.

K**kinase**

Enzyme that catalyzes the addition of phosphate groups to molecules.

Kramers-Kronig transform

A general formalism that relates the real part of any analytic complex function to its imaginary part. Useful for physical response functions such as an X-ray absorption spectrum.

***k*th raw moment** – see **raw moment**

kurtosis

The fourth **central moment** of a distribution; a measure how “squat” or “peaky” a distribution is.

L

Lambert–Beer absorption law (6-54)

Basic exponential law for the absorption of electromagnetic radiation.

lattice

An infinite, periodic 2-dimensional or 3-dimensional mathematical construct defined by three unit lattice vectors.

lattice parameters

The dimensions (norm, magnitude) a , b , c , of the **lattice vectors** \mathbf{a} , \mathbf{b} , \mathbf{c} , and the enclosed angles α , β , and γ . Identical to **unit cell parameters**.

lattice planes

Sets of planes in a real lattice $[0, \mathbf{a}, \mathbf{b}, \mathbf{c}]$ spanned by three **lattice points** with indices (ua, vb, wc) where u , v , w , are integers and known as direct **Weiss indices**. Lattice planes are more commonly described or indexed by reciprocal **Miller indices** h , k , l corresponding to a reciprocal lattice point $\mathbf{r}^* = (h\mathbf{a}^* + k\mathbf{b}^* + l\mathbf{c}^*)$.

lattice point

Any point $\mathbf{r} = (u\mathbf{a} + v\mathbf{b} + w\mathbf{c})$ with u , v , w integers in a lattice with basis $[0, \mathbf{a}, \mathbf{b}, \mathbf{c}]$. Centered lattices possess additional lattice points corresponding to their **Bravais translations**. Compare **reciprocal lattice point**.

lattice spacing

The shortest distance $d(hkl)$ (d_h) between **lattice planes** in direct space indexed with reciprocal **Miller indices** hkl . The interplanar distance vector \mathbf{d}_h is collinear with reciprocal lattice vector \mathbf{d}_h^* and reciprocal to it in **norm** or length. See also **resolution**.

lattice symbols

Symbols used for the 14 **Bravais lattices**. Table 6-6.

lattice types – see **Bravais lattices**

lattice vector

A vector from the origin (or other lattice point) to a **lattice point**.

Lattman angles

A special combination of **Eulerian angles** which prevents correlation problems at certain angle combinations.

Laue diffraction

Diffraction method that uses a broad **bandwidth** of X-ray energies instead of monochromatic X-rays. More reflections fulfill the diffraction condition at the same time as the **Ewald sphere** now becomes a Ewald shell with a thickness corresponding to the inverse of the bandwidth limits. Allows **time-resolved X-ray diffraction** experiments for the elucidation of enzyme mechanism and transport reaction.

Laue equations (6-20)

Set of three basic equations defining diffraction conditions in three dimensions.

Laue group (class)

Derived from the crystallographic **point groups** by addition of **inversion**, generating 12 distinct Laue groups (point group 32 splits into two Laue groups depending on the 2-fold axis orientation as defined by the corresponding **trigonal** space groups).

Laue symmetry

Point group symmetry plus **inversion**; the symmetry of **reciprocal space**. Important for efficient data collection strategies. Table 6-6.

LBFGS

A quasi-Newton optimization algorithm (limited-memory Broyden–Fletcher–Goldfarb–Shanno). LBFGS is particularly well suited for optimization problems with a large number of dimensions, as it does not explicitly form or store the Hessian (second derivative) matrix.

lead optimization

Improvement of desirable properties such as target binding and specificity of a drug lead compound by chemical modification.

least squares minimization – see **least squares refinement**

least squares refinement

General refinement method that minimizes the **sum of residuals squared** (SRS; hence the name) between observed and calculated data by adjusting model parameters. In crystallographic refinement, the squared differences between observed and calculated structure factor amplitudes are minimized by adjusting positional parameters (coordinates) and **B-factors** of the model atoms. Least squares methods are a special case of **maximum likelihood** methods applicable for complete models with only random errors.

Lennard-Jones potential

Potential energy curve of the basic form $Ar^{-12} - Br^{-6}$. Figure 12-18.

leucine

Hydrophobic L- α -amino acid with a branched aliphatic $-\text{CH}_2\text{CH}(\text{CH}_3)_2$ side chain.

leucine zipper

Structural motif in DNA-binding proteins in which two alpha helices from separate proteins are joined together in a coiled-coil (rather like a zipper), forming a protein dimer. Figure 2–47.

libration

From Latin *librare* “to balance, to sway.” Rocking motion of a group of atoms along an arc, used in the **TLS** parameterization of molecular motion.

ligand

Any molecule that binds to a specific site on a protein or other molecule.

ligand docking

Molecular modeling technique that finds and scores the orientation and conformation (pose) of a ligand molecule.

ligase

Enzyme that ligates two molecules in an energy-dependent process. DNA ligase joins two DNA molecules together end-to-end through phosphodiester bond formation.

light scattering methods

Non-destructive optical methods for conformational analysis of macromolecules. **Dynamic light scattering** and **static light scattering** methods provide information about hydrodynamic properties, conformational state, oligomerization state, and molecular mass.

likelihood

The probability of random variables in an experiment with known outcomes, that is, the estimate or adjustment of model parameters based on known outcomes. In other words, likelihood is the “inverse” of basic probability, which is the prediction of outcomes given the parameters. Compare **probability**.

likelihood by intuition

A number of chemically and physically reasonable common-sense assumptions incorporated in George Sheldrick’s successful (non-maximum likelihood) programs.

likelihood enhanced translation function (LETF)

Translation function that is faster than the full maximum likelihood translation function, used in initial stages of translation search.

likelihood function (L)

A probability function describing the probability of variables or parameters of a model given certain experimental outcomes or data.

likelihood ratio

Ratio of (posterior model) likelihoods of two competing hypotheses, based on the *same* data. See also **Bayes factor**.

linear merging R-value (8-15)

linear regression

Modeling or fitting of relationship between one or more independent variables and one dependent (response) variable by a linear **least squares** regression function. A linear regression with one independent variable represents a straight line. Linear regression does not imply a straight line, but refers to the linearity of the regression function.

linear residual, deviate (7-28)

The difference between two values, often between observed and calculated values or observed and **expectation values**.

linearly independent

A set of vectors (or equations) is linearly independent if none of them can be written as a linear combination of finitely many other vectors (or equations) in the set. The **determinant** of a matrix describing a linearly independent set of vectors (or set of equations) is nonzero.

lipid cubic phase

The lipid mono-olein forms a complex phase system with water. One of the phases in the mono-olein/water system is a bilayered, cubic phase containing 50–80% lipid as well as interconnected solvent channels. Bacterial rhodopsins, halorhodopsins, and photosynthetic reaction centers have been crystallized in lipid cubic phases.

liquid nitrogen, LN₂

Colorless clear **cryogenic** liquid with density of 0.807 g·ml⁻¹ at its boiling point of 77 K (–196°C; –321°F), produced by fractional distillation of liquid air. Used to cool **protein crystals** to cryogenic temperatures during X-ray exposure in order to minimize **radiation damage**.

local minimum

A parameter combination in a multi-dimensional solution landscape that is not the global (absolute) minimum.

local pK_a

Low **solvent polarity** raises the **pK_a** of an acid residue, because it destabilizes the ionized form. Conversely, if a basic lysine residue is buried in a hydrophobic environment, its **pK_a** will be lower.

local refinement – *see* **real space refinement****local scaling**

Scaling method where data sets are scaled together in local groups instead of simple overall **Wilson scaling**.

locked rotation function (11-34)

A self-rotation or cross-rotation function used for searching for multiple molecules, in which the NCS relation between the molecules is known and held fixed.

locked translation function

A **translation function** used for searching for multiple molecules, in which NCS between the molecules or already positioned molecules or parts thereof are accounted for.

log-likelihood (LL)

The natural log_e (ln) of a likelihood function *L*. Reduces multiplication of likelihood functions to sum of log-likelihood functions when forming **joint probabilities**.

log-likelihood gain (LLG)

The *LLG* is a measure for the improvement of the model during refinement or when comparing different models against the same data. In crystallographic model comparison, the *LLG* is commonly computed as the difference between the log-likelihood of the model and the log-likelihood calculated from a Wilson distribution, thus measuring how much better the data can be predicted with a particular model than with a random distribution of the same atoms.

loop (B)

Sequence of residues of varying length, connecting secondary structure elements, with no specific hydrogen binding pattern and thus not a secondary structure element. Often disordered in crystal structures. *Compare* **turn**.

Lorentz–Cauchy distribution

The form of a spectral line, similar—in shape only—to a **Gaussian normal distribution**, but with unpleasant properties such as undefined mean and infinite variance.

Lorentz exclusion region – *see* **Lorentz factor****Lorentz factor** (6-51)

A correction factor for raw diffraction intensities taking into account the different amount of time it takes for a **reciprocal lattice point** to pass through the **Ewald sphere**. The shallower (tangential) the intersection with the Ewald sphere, the longer the reciprocal lattice point remains in **reflection condition** during the recording of a rotation image. Points along the rotation axis can remain in reflection conditions too long to be properly corrected, and fall into the **Lorentz exclusion region**. Often combined with the **polarization factor**.

low energy remote data set

Anomalous diffraction data collected about 100 eV or so below the absorption edge. Sometimes used for additional dispersive signal relative to **inflection data set**, but carries no **anomalous signal** for the edge atom.

low resolution structure

Structure refined against data with resolution of less than ~3.5 Å (approximately the **determinacy limit** for restrained coordinate refinement, therefore **torsion angle** refinement with fewer adjustable parameters is commonly used at low resolution).

lunes

Elliptic regions of reflections recorded on an area detector during a rotation increment.

Luzzati D-factor

A factor by which the amplitude of a model structure factor is reduced due to positional (atom coordinate) and scattering factor errors.

Luzzati plot

A graph of *R*-value plotted versus $\sin \theta / \lambda$ where the slope in its linear region provides an upper estimate of the mean coordinate error.

lysine

Positively charged L- α -amino acid, long side chain $-(\text{CH}_2)_4\text{NH}_3^+$ with high surface entropy.

lysine methylation

Mild chemical methylation of surface lysine residues in proteins, aiming to improve crystallization.

lysozyme

Enzyme that catalyzes the cutting of polysaccharide chains in the cell walls of bacteria. A hardy perennial of protein crystallization, for practice and physical chemistry studies, with atypically high **supersaturation** levels.

M**macromolecular refinement**

Conducted as restrained global reciprocal space refinement against experimental structure factor amplitudes or local real space refinement against electron density.

macroscopic twinning

Intergrowth of crystal in multiple regions or domains that are highly misaligned and generally can be visually distinguished and sometimes mechanically separated.

MAD – *see* **multi-wavelength anomalous diffraction phasing****magic angle**

Angle between the base of a cube and its space diagonal of 54.74°, ($\tan^{-1} \sqrt{2}$).

magnitude (M)

Length or **norm** of a vector; amplitude of the electric field vector.

main chain atoms

The atoms N, C α , C, and the carbonyl oxygen O of each residue form the main chain (backbone) of a polypeptide.

main chain torsion angle plot

Representation of pairs of ϕ - ψ torsion angles of each residue in an energy contour plot, representing a potential energy surface. Repulsive **van der Waals interactions** limit the probable torsion angles to certain regions, typical for the **secondary structure** in which the residues partake. As the main chain torsion angles are generally not restrained during refinement, the torsion angle analysis provides valuable stereochemical **cross-validation**. Also known as Ramachandran plot, named after its inventor.

main chain torsion angles

The **torsion angles** ϕ (ϕ), ψ (ψ), and ω (ω) around N and C α of a residue, C α and C of a residue, and around C of one residue and N of the subsequent residue in a polypeptide chain, respectively.

map

A graphical representation of a 2-dimensional or 3-dimensional **function** where points of equal value are connected with 2-dimensional contour lines or a 3-dimensional mesh or grid. *See also* **electron density map**, **Patterson map**.

map averaging

Procedure of averaging electron density, often combined with **solvent modification** for phase improvement.

map contrast and connectivity

Measures for electron density map quality. Contrast between protein region and solvent region, as well as connectivity within the protein region, should be high.

map inversion

Computation by Fourier transformation of complex structure factors from electron density.

map-likelihood phasing

Procedure of density modification based on the adjustment of the initial electron density map likelihood toward the likelihood function of a proper protein map.

marginalization

Elimination of a **nuisance variable** from a **probability distribution** by integrating it out.

marker atom

Atoms incorporated in a protein structure or crystal providing the source of **isomorphous differences** or **anomalous differences** in diffraction data. Common markers are **heavy metal** ions, native or soaked/co-crystallized with proteins; Se in **selenomethionine** labeled proteins; or native sulfur in case of **S-SAD phasing**.

marker atom substructure – see substructure**masked bulk solvent model – see flat bulk solvent model****mass absorption coefficient**

Measure for how much X-rays are absorbed by a certain element, in $\text{cm}^2 \text{g}^{-1}$.

mass spectrometry (MS)

Technique for identifying compounds on the basis of their mass-to-charge ratio. Common uses are identification of proteins, determination of molecular mass, sequencing polypeptides, or hydrogen–deuterium exchange mass spectroscopy.

mathematical group – see group**matrix operators**

Representation of a symmetry operation as a combined rotation and translation matrix. *See also* **symbolic operators**.

matrix weight

Term used by *REFMAC* for the **weight** of the X-ray term, w_a in *CNS* terminology.

Matthews coefficient

Ratio of asymmetric unit volume over protein molecular weight, on average $2.5 \text{ \AA}^3/\text{Da}^{-1}$; related to solvent content (11-1).

Matthews probability

Conditional probability that the **asymmetric unit** of a crystal

structure harbors a homooligomer, given the diffraction limit or **resolution** of the data.

maximum entropy method

Special method of direct phase determination based on maximum (informational) entropy distributions of atoms in a structure, can use prior information.

maximum likelihood – see maximum likelihood principle**maximum likelihood coefficients**

Fourier coefficients for electron density reconstruction derived from **maximum likelihood functions** based on **sigma-A**.

maximum likelihood function

A joint probability distribution that can accommodate incompleteness and various errors (as well as prior information) of the model. *Compare* **maximum posterior refinement**.

maximum likelihood method

Any optimization method that maximizes the **data likelihood function** in order to obtain the proportional **model likelihood**.

maximum likelihood principle

States that the best estimate for any parameter is the value that maximizes a **likelihood function**, that is, the joint probability of experimental outcomes given the model parameters (as well as prior information).

maximum likelihood rotation function MLRF (11-57)

Rotation function based on conditional Rice and Woolfson distributions accounting for relative model phases and data errors.

maximum likelihood translation function MLTF (11-53)

Translation function based on Rice and Woolfson distributions, accounting for model incompleteness and errors.

maximum posterior refinement

A maximum likelihood method that incorporates prior probability (commonly in the form of stereochemical restraints) in the maximum likelihood target function. Sometimes called maximum *a posteriori* refinement or regularized maximum likelihood refinement.

mean (M)

Expectation value of a **probability distribution**, in simple cases the arithmetic average or a weighted average.

mean absolute error (7-38)**mean E-squared minus one ($\langle |E^2 - 1| \rangle$)**

Mean absolute of normalized structure factor squared minus one, indicating whether a structure is centrosymmetric (0.968) or acentric (0.736). Can be less for twinned structures.

median (M)

The median of a **probability distribution** $P(x)$ is the value of x where the integral over the distribution (area) above and below x are equal. *Compare* **mean**.

membrane

The fluid lipid bilayer plus associated proteins that encloses all cells and, in eukaryotic cells, many organelles as well.

membrane protein

Protein associated or anchored with, or extending across, lipid membrane of cells. Commonly used for transmembrane proteins that have one or more **transmembrane helices**. *See also* **integral membrane protein**.

merging (C)

Combining various measurements of identical or symmetry related reflections into one (unique) data set.

merging R-value – see linear merging R-value (8-15)**merohedral twinning**

Growth of domains in different orientations that can be precisely described by a symmetry operation compatible with, or extending, the crystal symmetry. Reflections from the distinctly oriented **domains** perfectly superimpose and the diffraction pattern looks unsuspecting and normal. The symmetry operator is called the

- twin operator.** Merohedral twinning can be detected by deviations from expected **intensity distributions**. *See also hemihedral twinning.*
- messenger RNA (mRNA)**
RNA molecule that specifies the amino acid sequence of a protein. Produced in eukaryotes by processing of an RNA molecule made by RNA polymerase as a complementary copy of DNA. It is translated into protein in the translation process catalyzed by **ribosomes**.
- metabolic inhibition**
Suppression of Met biosynthesis through inhibition of aspartokinase by excess Ile, Lys, and Thr in the medium. Used to overexpress **selenomethionine** labeled proteins.
- metastable**
State of a system where upon disturbance or **nucleation** it returns under phase separation to equilibrium.
- methionine**
A hydrophobic amino acid with a $-\text{CH}_2\text{CH}_2\text{SCH}_3$ side chain. *See also seleno-methionine.*
- methyl group**
Hydrophobic $-\text{CH}_3$ group.
- methylene group**
Hydrophobic $-\text{CH}_2-$ group.
- metric tensor, G**
Dyad product of the basis vectors; used in molecular geometry calculations.
- Metropolis–Monte Carlo algorithm**
Stochastic optimization procedure based on random parameter changes; accepts a certain fraction of uphill movements.
- microfluidics**
The precise control and manipulation of fluids that are geometrically constrained to a small, typically sub-millimeter, scale. Microfluidic chips made of transparent silicone rubber are used in protein crystallization experiments and require very little material.
- microseeding**
Introduction of microscopic, external **nucleation** seeds into crystallization drop by **streak seeding** or other methods.
- Mie ratio**
The ratio of particle radius over **wavelength** determining the characteristics of a scattering process. *See also Rayleigh scattering.*
- Miller indices**
Integer number triple h, k, l written as reciprocal index vector **h** specifying sets of parallel, equidistant 3-dimensional **lattice planes** (Figure 5-46). Derived from direct **Weiss indices** by inversion and normalization with smallest common denominator.
- MIR** – *see multiple isomorphous replacement*
- MIRAS**
Multiple isomorphous replacement with anomalous signal, uses orthogonal anomalous phasing to improve **phase probabilities**.
- mirror operation**
A **symmetry operation** that generates a mirror image of an object. Not allowed for **chiral** motifs such as protein molecules, because mirror operations change the **handedness** of the motif.
- mitogen, mitogenic**
Extracellular signal molecule that stimulates cells to proliferate.
- mode (M)**
Peak value of a probability distribution, most probable value. The **mean** and the **mode** are equal only in monomodal symmetric distributions.
- model (C)**
The physical model of atoms representing the crystal structure. The same physical model can be differently parameterized, that is, represented by a different *mathematical model*.
- model (M)**
The mathematical set of adjustable parameters, describing a *physical* model.
- model bias**
Tendency of calculated **phases** to dominate the electron density reconstruction, notably in case of **molecular replacement** phasing.
- model building (C)**
Process of manually or by means of automated programs placing the atomic model of the protein structure into the **electron density**.
- model likelihood (function)** – *see posterior probability*
- modulated structure (C)**
Crystal structure with additional periodicity superimposed over crystal lattice period. Can be **commensurate** or **incommensurate** with the lattice period.
- molecular dynamics (MD)**
Potential energy minimization based on empirical potential parameterization and minimization of force equations. Can include X-ray terms or NOEs for experimental structure refinement.
- molecular envelope** – *see scattering envelope*
- molecular replacement**
Phasing method that uses the calculated phases from a correctly placed, structurally similar model to phase an unknown structure. Commonly broken up into a 3-dimensional rotational and subsequent 3-dimensional translational search. Can be based on 6-dimensional **stochastic searches**, **Patterson search** methods, and modern **maximum likelihood**-based search functions.
- monochromator**
Selects a narrow energy band of X-rays. Commonly one or more crystal monochromators separating X-rays by Bragg reflection. Focusing multi-layer X-ray mirrors can replace/enhance monochromators.
- monoclinic**
Lattice or crystal structure with cell parameters $a \neq b \neq c$, $\alpha = \gamma = 90^\circ$, $\beta \neq 90^\circ$, with primitive (*mP*) and *C*-centered (*mC*) lattice types. Minimal internal symmetry a 2-fold axis parallel to **unique axis b**.
- monoclonal antibody**
Antibody secreted by a **hybridoma** cell line. Because the hybridoma is generated by the fusion of a single B cell with a single tumor cell, each hybridoma produces **antibodies** that are all identical.
- monomer (B)**
Single molecular entity, component of **oligomer**.
- monomer (C)**
Molecular entity for which a **stereochemical restraint** file exists in a restraint file library (**CCP4**).
- monotopic membrane protein**
Protein associated or anchored in the lipid membrane of a cell with a simple tail or membrane anchor. *Compare bitopic membrane protein.*
- morphology (C)**
Shape and form of a crystal. *See also polymorphism.*
- mosaic crystal**
Single crystal consisting of multiple, slightly misaligned blocks or **domains**.
- mosaicity**
The degree of misalignment between **domains** of a **mosaic crystal**.
- mother liquor**
The liquid medium surrounding a growing protein crystal and filling its **solvent channels**.
- motif (B)**
Element of structure or pattern that recurs in many contexts, particularly a small structural pattern, often with specific function, that can be recognized in a variety of proteins.
- motif (C)**
Physical contents of the asymmetric unit which makes up the physical crystal structure upon application of **space group symmetry operations**. Can be atoms, parts of a molecule, an entire molecule, or an **oligomer**.

mtz file

Binary data format used by the CCP4 suite and other programs.

multi-conformer refinement

Refinement protocol that captures the dynamic nature of molecules by refining multiple models using time-averaged molecular dynamics.

multimodal

A function or probability distribution with multiple peaks or local extrema.

multiple conformations

Two (or rarely more) discrete conformations of the same molecule or part thereof, often observed in case of surface-exposed residues or ligand molecules.

multiple isomorphous replacement (MIR)

Method of **isomorphous replacement** which uses several heavy atom derivatives to resolve the **phase ambiguity**.

multiplicity of space group (C)

Multiplicity of the **point group** times the multiplicity of the **Bra-vaïs translation** (and times 2 in case of a **centrosymmetric** space group). Equal to the number of **asymmetric units** in the unit cell.

multi-solution methods

Methods for (sub)structure solution that generate and rank multiple possible solutions.

multivariate

A function or probability distribution depending on multiple variables or parameters.

multi-wavelength anomalous diffraction phasing (MAD)

Anomalous phasing technique utilizing both **anomalous** and **dispersive differences**.

mutagen, mutagenic

A mutagen is a physical or chemical agent that changes the genetic material of an organism, generating mutations. As many mutations cause cancer, mutagens are typically also carcinogens.

N**N-terminus – see amino terminus****native Patterson map**

A map of the Patterson function computed from the intensities of a native protein structure as Fourier coefficients, compared to a difference Patterson map.

NCS – see non-crystallographic symmetry**neutron**

Subatomic particle with no net electric charge carrying a magnetic moment and a mass slightly larger than that of a **proton**. Neutrons are produced in nuclear fission reactors or **spallation neutron sources**. For **neutron diffraction** they are moderated (thermal neutrons) so that their energy is equivalent to a wavelength on the order of 1 Å. Kinetic energy, speed, and **wavelength** of the neutron are related through the **de Broglie relation**.

neutron diffraction

Technique using thermal **neutrons** instead of X-rays for diffraction on crystals. Advantage is that the neutron scattering factor for hydrogen and deuterium is different and of the same order as for heavy atoms, therefore hydrogen atoms can be located and hydrogen–deuterium exchange studied.

nitrogen

Colorless, odorless, tasteless, and mostly inert diatomic gas at standard conditions, constituting 78% by volume of the earth's atmosphere. *See also* **liquid nitrogen**.

NMR – see nuclear magnetic resonance spectroscopy**NOESY**

Nuclear Overhauser effect (NOE) spectroscopy, 2-dimensional NMR technique for structure determination of macromolecular structures. NOE resonances are observed through space, not through bonds. *Compare* **COSY**.

non-centric reflections – see acentric reflections**non-commutativity (M)**

Outcome of operation is dependent on the order of elements, that is, $ab \neq ba$.

non-crystallographic symmetry (NCS)

Presence of more than one copy of a motif in the **asymmetric unit**. Frequently the case; nearly half of all structures in the PDB are **dimers** or higher **oligomers** state of the molecules. Molecules may be related through a general combination of rotation and translation; NCS is local and not limited to crystallographic **symmetry operations**. *See also* **proper non-crystallographic symmetry**, **translational non-crystallographic symmetry**. Figures 5-19, 11-2, 11-12.

non-merohedral twinning

Non-merohedral twinning can occur if cell dimensions of a crystal are such that they either *match* or are *reasonably close* in two dimensions when domains are specifically oriented to each other. In this case, the crystals look visually unsuspecting, but the diffraction pattern in three dimensions will reveal two different interpenetrating lattices. If the spots are reasonably resolved and one orientation dominates, the major component can often be indexed and integrated separately.

norm (M)

Euclidean norm: length, magnitude, of a vector, square root of sum of squared vector components.

normal distribution – see Gaussian normal distribution**normal equations**

An $n \times p$ system of linear equations used in the setup of least squares minimization, with n the number of observations and p the number of model parameters.

normalized structure factor distribution

Probability distribution functions for **normalized structure factor** (or **structure factor amplitude**) to have a certain value. Can be unconditional (centric and acentric **Wilson distribution**), or conditional on a model (centric **Wolfson**, acentric **Sim** or **Rice distribution**).

normalized structure factors (7-109)

Structure factor amplitudes E_h reflecting the scattering of point atoms, adjusted for **B-factor** attenuation, **atomic scattering factor** attenuation, and **epsilon-factor**. *Compare* **unitary structure factors**.

normalized sum of residuals squared – see chi-square**nth central moment – see central moment****nuclear magnetic resonance spectroscopy (NMR)**

Spectroscopic technique for solution structure determination that exploits the resonance of magnetic states of spin-carrying nuclei induced by radio frequency **electromagnetic radiation**.

nucleation (C)

Formation of sites from which phase separation proceeds, either due to homogeneous formation of nucleation sites at high supersaturation or heterogeneous introduction of nuclei (e.g. foreign matter or nucleation **seeding**).

nucleic acid

RNA or DNA, a macromolecule consisting of a chain of nucleotides joined together by phosphodiester bonds.

nucleophile

A **nucleophile** (literally “nucleus lover”) is a chemical entity that seeks to bind to its reaction partner (the electrophile) by donating bonding electrons.

nucleoside

Purine or pyrimidine base covalently linked to a ribose or deoxyribose sugar.

nucleosome

Beadlike structure in eukaryotic chromatin, composed of a short

length of DNA wrapped around an octameric core of histone proteins. The fundamental structural unit of chromatin.

nucleotide

Nucleoside with one or more phosphate groups joined in ester linkages to the sugar moiety. DNA and RNA are polymers of nucleotides. Figure 2-43.

nucleotide excision repair

DNA repair pathway in which entire nucleotides are removed from the DNA helix and replaced. *Compare* **base excision repair**.

nuisance variable

A parameter whose value is generally not known but which appears in a **probability distribution**. It can often be eliminated from the probability distribution by integrating it out (**marginalization**).

Nyquist theorem

If a discrete function $F(x)$ contains no terms higher than N (corresponding to $1/d_{min}$) its transform is completely determined by representing it by a series of points spaced $1/2N$ (that is $d_{min}/2$) apart.

O

oblique (C)

A lattice or crystal system that is not **orthogonal**, also specifically identifying a 2-dimensional plane lattice with $a \neq b$, $\gamma \neq 90^\circ$ or 120° .

observations (C)

Any experimental **data** point such as measured data or knowledge-based **restraints**. In restrained refinement, experimental data (structure factor amplitudes or intensities) as well as independently established model restraints.

obverse (C)

Definition of normal or standard assignment of crystallographic axes; primitive **rhombohedral** cells are indexed as rhombohedrally centered **hexagonal** cells in obverse setting when $R = (2/3 \ 1/3 \ 1/3)$, $(1/3 \ 2/3 \ 2/3)$.

occupancy – see occupancy factor

occupancy factor

Fraction $0 \leq n \leq 1$ of atoms or molecules that actually occupy a crystallographic position, often less than one for ligands, depending on the **dissociation constant**.

occupancy refinement

Used in substructure solution to improve the solution. A sharp drop in occupancy after the expected number of marker atoms is indicative of a correct solution.

oligomer

Complex made of a number of subunits. *See also* **heterooligomer**, **homooligomer**.

oligosaccharide

Short linear or branched chain of covalently linked sugars.

omega backbone torsion angle

The backbone **torsion angle** defining the torsion around the C atom of one **residue** and N atom of the subsequent residue in a polypeptide chain. Restrained to 180° . Figure 2-7.

omega-loop – see loop

omit map

Electron density reconstruction from **Fourier coefficients** in whose computation a questionable part of the model has been omitted. Used as means of reducing **model bias**, optionally in combination with coordinate perturbation and/or **simulated annealing**. To generate composite omit maps, a different block of model is omitted each time the map is calculated, and the final composite map is averaged from the individual maps.

optical transforms

Fourier transforms of objects obtained experimentally by

diffraction with visible light, used as a means of trial-and-error structure determination in the very early days of crystallography, and still useful for educational purposes.

optimization algorithm

Optimization algorithms are procedures that search for an optimum of a nonlinear, multi-parametric function. Deterministic optimizations such as the gradient-based **maximum likelihood methods** are fast and work well when we are reasonably close to a correct model, at the price of becoming trapped in local minima. Stochastic procedures employ a random search that also allows movements away from local minima. They are slow but compensate for it with a large convergence **radius**.

optimization experiment

A crystallization experiment that follows initial crystallization trials, with the objective of narrowing down the parameter space and adjusting initial conditions to obtain improved crystals.

origin (C)

The point from which the three **unit cell vectors** in a crystal structure can originate. Many space groups allow multiple origins.

orthogonal

A system whose **basis** vectors are perpendicular to each other.

orthogonalization matrix

Describes a 3×3 similarity matrix used to transform a **crystallographic coordinate basis** into an **orthogonal Cartesian basis**. Its inverse is the deorthogonalization matrix.

orthologs

Genes or proteins from different species that are similar in sequence because they are descendants of the same gene in the last common ancestor of those species. *Compare* **paralogs**.

orthonormal

An **orthogonal** system whose basis are **unit vectors**.

orthorhombic

Lattice or crystal structure with cell parameters $a \neq b \neq c$, $\alpha = \beta = \gamma = 90^\circ$, with primitive (*oP*), *F*-centered (*oF*), or *I*-centered (*oI*) lattice types. Minimal symmetry three non-intersecting 2-fold rotation axes.

oscillation method – obsolete; see rotation method

outlier detection

Statistical method to flag data that exceed a certain deviation from the expectation value.

overall B-factor

Isotropic overall *B*-factor, obtained by **Wilson scaling** of diffraction data. During subsequent refinement, generally replaced by **anisotropic** overall displacement tensor.

overexpression

High expression levels of a protein in the native host or in other organism (heterologous expression).

overfitting

Introduction of too many parameters into mathematical refinement model, often by adding too many waters or other obscure features into the structure model. Controlled by **cross-validation**.

P

p53

Tumor suppressor gene found mutated in about half of human cancers. Encodes a gene regulatory protein that is activated by damage to DNA and is involved in blocking further progression through the cell cycle. Figure 2-47.

palindromic sequence

Nucleotide sequence that is identical to its complementary

strand when each is read in the same chemical direction, for example GATC.

paralogs

Genes or proteins that are similar in sequence because they are the result of a gene duplication event occurring in an ancestral organism. *Compare orthologs.*

parameter

Variable in a mathematical model.

parameter estimation (M)

Process of estimating parameters of a model based on experimental data. An estimator takes data as input and provides model parameters as output. Common estimators in crystallography are least squares, maximum likelihood, or maximum *a posteriori* estimators.

Parseval's theorem

States that the mean square of a complex function $F(\mathbf{x})$ and the mean square of its Fourier transform are proportional. Sidebar 9-5.

partial derivative

Derivative of a multi-parametric function with respect to one specific parameter.

partial occupancy – see occupancy factor

partial reflection, partial

An X-ray deflection or reciprocal lattice point extending over multiple diffraction images or frames is a “partial” on each frame.

partial structure (C)

Molecular structure where a part is unknown, accounted for with **Sim weights**, or if model errors are also included, with the **sigma-A** parameter.

Patterson correlation function

A scoring function based on the correlation coefficient between two Patterson maps.

Patterson correlation refinement

Refinement of atom or molecule positions by optimizing the correlation between observed and calculated Patterson functions. Requires only correct orientation, translation invariant.

Patterson cross-rotation (function)

Used in **molecular replacement** Patterson rotation searches, based on the overlap of **cross-Patterson vectors**.

Patterson cross-vectors

Interatomic distance vectors between different atoms in the asymmetric unit in a Patterson map. Not the same as **cross-Patterson vectors**.

Patterson function (9-29)

The Patterson function is the **autocorrelation** of the **electron density**, which is equivalent to the Fourier transform of the squared **structure factor amplitudes**, or the reflection intensities. The Patterson function has its maxima at the interatomic distance vectors between all atoms in the unit cell. It is inherently **centrosymmetric** and its symmetry is the **space group** symmetry plus **inversion**. It is used for (i) **substructure** solution, (ii) **molecular replacement** searches, and (iii) detection of translational NCS parallel to unit cell axes and proper NCS axes parallel to crystallographic axes via **native Patterson maps**.

Patterson map

A contoured 2- or 3-dimensional representation of the **Patterson function**.

Patterson minimum function

A robust scoring function for the correctness of a trial substructure; used in *SHELXD*.

Patterson search

A search algorithm that is based on the **Patterson function**, matching **interatomic distance vectors** of the search probe and those obtained from experimental data. Can be a **Patterson self-rotation** search, a **Patterson cross-rotation** search, or a Patterson translation search.

Patterson seeding

Method to pick starting phases (atoms) for **direct methods** from Patterson superposition, implemented in substructure solution program *SHELXD*.

Patterson self-rotation (function)

Used in Patterson searches, based on the overlap of **self-Patterson vectors**. *See also Patterson self-rotation map.*

Patterson self-rotation map

A map containing in **stereographic projection** the peaks of the **Patterson self-rotation function**, which reveal the orientation of proper non-crystallographic n -fold symmetry axes in the $\kappa = 360/n$ degree section. Figure 11-12.

Patterson self-vectors

Interatomic distance vectors between the same atom and its symmetry mates in a Patterson map. Not the same as **self-Patterson vectors**.

Patterson superposition methods – see Patterson vector methods

Patterson symmetry

Space group symmetry plus inversion.

Patterson vector methods

Patterson function-based search method that uses atom pairs or trial structures to locate (heavy) atom positions.

PCR – see polymerase chain reaction

peak data set

Anomalous diffraction data collected at the absorption edge maximum defined by an **absorption edge** scan.

PEG – see poly(ethylene glycol)

PEG-MME – see poly(ethylene glycol) monomethyl ether

pegylation

Attachment of drug molecules or enzymes to PEG for better delivery and release properties.

peptide – see polypeptide

peptide bond

Chemical bond between the **carbonyl group** of one **amino acid** and the **amino group** of a second amino acid—a special form of amide linkage. Peptide bonds link amino acids into long macromolecular **protein** chains. Almost always in **trans-conformation**, only one in 1000 peptide bonds assumes the less favorable **cis-conformation** for non-**proline** residues.

perfect hemihedry – see hemihedral twinning

phase ambiguity

The phasing equations based on a single substructure always have two solutions for the phase angle. The ambiguity can be broken by (i) additional independent substructures (isomorphous derivatives), (ii) additional orthogonal **anomalous** data, or (iii) **density modification** techniques.

phase angle

The angle by which the phase of a **complex structure factor** or reflection is shifted from a reference phase, commonly zero. Readily computed for known structures, but requires **experimental phasing** for unknown structures.

phase bias – see model bias

phase combination (10-41)

Weighted averaging of experimental and/or model phase probabilities from different sources, yielding best phases from **joint phase probabilities** with **phase probability distributions** frequently expressed using **Hendrickson–Lattman coefficients**.

phase diagram

A diagram depicting the phase relations in a multi-component system. *See also crystallization diagram.*

phase extension

Procedure which provides phases for (high resolution) reflections, for which no initial experimental phases from the marker

atom substructure solution are available. Solvent flattening, solvent flipping, sphere of influence, histogram matching and, Free Lunch are examples for common phase improvement and phase extension methods.

phase probability distributions

The probability that an experimentally determined phase angle has at each angle in the range 0 to 2π rad. Determined by solving the phasing equations, often in the form of specific **maximum likelihood functions**.

phase problem

As the intensity of diffracted X-rays is given by the product of the **complex structure factor F** and its **complex conjugate F***, it is **Hermitian** and the result is thus real, as required for a physical observable. This means that phase information is lost during the measurement process, creating the phase problem. The phases must be separately supplied by a variety of **phasing methods**. See also **experimental phasing**.

phase restrained refinement

Reciprocal space refinement that includes the independent experimental phase information as a restraint for the structure factors in the likelihood target function.

phase restricted reflections

Centric reflections are phase restricted.

phase separation

Occurs when a **metastable, supersaturated** system returns to equilibrium.

phased translation function

Translation function that uses electron density maps and thus phase information to determine translation vectors of properly oriented maps or models.

phases – see phase angle

phasing

Colloquial term for establishing the **phase angles** associated with measured **structure factor amplitudes**. Necessary procedure to solve the crystallographic **phase problem**. See also **experimental phasing**.

phasing circles

Circles in a **Harker diagram**.

phasing marker – see marker atom

phasing methods

See Table 10-1

phasing power (10-58)

An occasionally found statistic defined for each derivative as the sum of heavy atom or anomalous contributions divided by the sum of **closure residuals**.

phenylalanine

Hydrophobic, aromatic L- α -amino acid with an aromatic $-\text{CH}_2\text{C}_6\text{H}_5$ side chain.

phi backbone torsion angle

The backbone **torsion angle** for the torsion around N and C α of a **residue**. Figure 2-7.

phi-psi torsion angle plot – see main chain torsion angle plot

phospholipid

Most common category of lipids used to construct biomembranes. Generally composed of two fatty acid tails linked through glycerol (or sphingosine) phosphate to one of a variety of polar head groups.

phosphorescence

Photoluminescence process where the absorbed energy generates a state of higher spin multiplicity. Once the energy is trapped in the triplet state, transition back to the lower singlet energy states is forbidden and thus slow, sometimes taking hours. See also **imaging plate**.

photon

Uncharged particle with no rest mass, of energy $E = h\nu$ with h the

Planck constant and ν the frequency. See also **energy-wavelength conversion**.

photon flux

Number of **photons** per second that pass through an area of 1 mm^2 $\Phi_\lambda = \text{photons}\cdot\text{s}^{-1}\cdot\text{mm}^{-2}$. Dimensions depend on the energy unit selected for the photon.

physiological pH

Commonly the pH of blood, $\sim \text{pH } 7.4$ Varies for different cellular environments.

pI – see isoelectric point

pi-helix

A rare, loosely packed helical secondary structure element with backbone hydrogen bonds from residue n to $n + 5$.

pi-stacking

An uncharged dipole interaction between π -electron systems of aromatic or **arginine** residues.

pK_a value

The negative logarithm of the acid dissociation constant, $-\log_{10}(K_a)$, with $K_a = [\text{H}^+]\cdot[\text{A}^-]/[\text{HA}]$.

plain rotation

A rotation (axis, operation) that does not possess any additional symmetry elements such as translation (in **screw axes**) or inversion (in **inversion axes**).

plane equation (M)

A plane can be defined by a point with coordinate vector \mathbf{x} on the plane and the normal vector \mathbf{h} to the plane, $\mathbf{h}\cdot\mathbf{x} = n$.

plane group (C)

Set of **symmetry operations** forming a mathematical **group**, giving rise to the general 17 (five chiral) different arrangements of motifs in a plane (2-dimensional) periodic structure.

plasmid

Small circular extrachromosomal DNA molecule that replicates independently of the genome. Modified plasmids are used extensively as vectors for DNA cloning.

pocketome

The pocketome of an organism is the collection of all possible small molecule binding envelopes present in its cells.

point group

Mathematical **group** of operations that keep the origin fixed at one point and comply with the rules for mathematical groups. See also **crystallographic point groups**.

Poisson distribution (7-20)

A one-parametric, discrete **probability distribution function** describing for example random counting errors.

polar

In the electrical sense, describes a structure (for example, a chemical bond, chemical group, molecule, solvent) with positive charge concentrated toward one end and negative charge toward the other as a result of anisotropic electron distribution. Polar molecules are likely to be soluble in water. See also **dielectric constant**.

polar coordinates

Define any point in a plane through a distance from origin (or pole) and the azimuth (or polar) angle.

polar solvent

Solvent consisting of polar molecules.

polar space group

Space group in which the origin is arbitrarily located on a crystallographic axis. Table 6-6.

polarizability

Proportionality constant quantifying the interaction of the electric field vector and the resulting induced electric dipole moment.

polarization (C)

Describes the orientation of the *electric field vector* relative to the propagation direction of the wave. The oscillations may be fixed in

a single direction (linear polarization), or the oscillation direction may rotate as the wave travels (circular or elliptical polarization).

polarization factor (6-50)

A correction factor for raw diffraction intensities taking into account the effects of instrument- and diffraction-geometry dependent **polarization** of scattered X-rays. Often combined with the **Lorentz factor**.

polychromatic

Radiation that contains photons with widely varying energy, also termed “white radiation.”

polycrystalline

Material that contains a randomly oriented sample of many small microcrystals.

poly(ethylene glycol) (PEG)

A common precipitant for protein crystallization, reducing protein **solubility** by competing for solvent water molecules (excluded volume effect). PEGs are organic polyalcohols HO-(O-CH₂-CH₂)_n-OH of varying chain lengths between ~200 and ~15 000 Da average molecular weight. PEGs above 1 kDa are solid white powders, and freshly prepared aqueous solutions are used as precipitants.

poly(ethylene glycol) monomethyl ether (PEG-MME)

Same as **poly(ethylene glycol)**, except PEG-MMEs have a CH₃O-(methoxy) group on one terminal CH₃O-(O-CH₂-CH₂)_n-OH.

polyion, polyionic

Molecule that carries multiple charges; proteins are polyionic macromolecules or polyionic **heteropolymers**.

polymer

Macromolecule made of many monomers.

polymerase

Enzyme that catalyzes polymerization reactions such as the synthesis of DNA and RNA.

polymerase chain reaction

Technique for amplifying specific regions of DNA by the use of sequence-specific **primers** and multiple cycles of DNA synthesis, each cycle being followed by a brief heat treatment to separate complementary strands.

polymorphism (B)

Refers to genes with two or more alleles that coexist at high frequency in a population.

polymorphism (C)

Refers to different **crystal forms** (morphologies) of the same material. *Compare habit*.

polypeptide

A linear **heteropolymer** built by combination of different proteino-genic **amino acids**. The term peptide is generally applied to shorter stretches or molecules, while larger polypeptides (> ~50 residues) folding into distinct **tertiary structures** are called **proteins**.

polyproline-II helix

A left-handed, helical secondary structure element consisting of *trans*-proline or *trans*-proline-rich stretches with backbone hydrogen bonds from residue *n* to *n* + 2. Long PP-II helices containing enzymatically C γ -hydroxylated prolines are building blocks of the twisted collagen fibers.

polytopic membrane protein

Integral membrane protein passing through the lipid membrane of a cell with multiple transmembrane helices and having multiple functional domains, commonly in the ectodomain, which sometimes can be expressed separately. *Compare monotopic membrane protein*.

pose (of a ligand)

The orientation of the ligand relative to the receptor including the specific conformation of the ligand when bound to the receptor.

posterior probability

Joint probability function describing the **model likelihood** $prob(model|data,I)$ as the product of **data likelihood function**

$prob(data|model,I)$ and **prior probability** $prob(model|I)$. *See also Bayes' theorem*.

posttranslational modification (PTM)

Covalent modifications to the expressed protein, commonly carried out by other enzymes. **Glycosylation**, N-terminal acetylation, phosphorylation, and myristylation are common PTMs.

powder diffraction

Diffraction method applied to polycrystalline material. Powerful for chemical characterization, limited in case of *de novo* structure determination.

precession camera

An ingenious, once popular, mechanical contraption invented by Martin Buerger that provides an undistorted projection of the reciprocal lattice onto film or detector.

precipitant

A primary component of the **crystallization cocktail** that reduces the protein solubility.

precipitate (C)

Describes, in the context of protein crystallization, a reagent that reduces the **solubility** of proteins.

precision

High precision indicates that measurements of a quantity have a narrow distribution (small variance) about their mean. An experiment can be precise but have low **accuracy**. Figure 7-4.

precision-indicating merging R-value (8-17)

preconditioned conjugate gradient minimization

Iterative multivariate minimization algorithm that approximates the second derivatives to improve the gradient descent path.

primary structure

The linear sequence of the amino acid **residues** in the protein chain, beginning at the **amino-terminus**.

primer

Oligonucleotide that pairs with a template DNA strand and promotes the synthesis of a new complementary strand by a polymerase.

primitive (C)

Non-centered **lattice**, **space group**, **crystal structure**; the six non-centered primitive **Bravais lattices** are *aP*, *mP*, *oP*, *tP*, *hP*, and *cP*. Primitive does not mean simple.

principal Euler angle

The matrix equation represented by the **direction cosine matrix R** describing a rotation can be solved and yields the three **eigenvalues**. The **eigenvector** belonging to the real eigenvalue 1 is the **Euler axis**, with the principal Euler angle defining the single rotation angle around the Euler axis **E**. Equivalent to the rotation angle κ in **spherical coordinate** representation.

prior probability

Probability distribution function describing the probability of a hypothesis or model prior to considering the specific experimental data of the present experiment. *See also Bayes' theorem*.

probability

The probability of a random variable is its *relative frequency of occurrence*, that is, it allows us to predict unknown outcomes based on known parameters. In general terms, also the **likelihood** of a certain outcome of an experiment, that is, the estimate of unknown parameters based on known outcomes. *Compare likelihood*.

probability density function

A continuous, integrable **probability distribution function**.

probability distribution function

General term for the probability of occurrence $P(x)$ of a random variable for given values of x .

projection (M)

Linear transformation described by matrix **P** fulfilling $\mathbf{P} = \mathbf{P}^2$. Used for example to project 3-dimensional objects (molecules or maps) onto a 2-dimensional plane.

prokaryote (procaryote)

Single-celled microorganism whose cells lack a well-defined, membrane-enclosed nucleus, either a bacterium or archaeon. Prokaryotes generally lack the capability of **posttranslational modifications**, have no dedicated translocation machinery, and their reducing cytoplasm often prevents disulfide formation when overexpressing eukaryotic (mammalian) proteins in prokaryotic hosts.

proline

Cyclic, hydrophobic, proteinogenic L- α -amino acid. The imide bonds between a non-proline residue and a subsequent proline are more frequently (~6%) found in *cis*-conformation than for two non-proline residues (0.1%). For the cyclic amino acid proline, the *cis*- and *trans*-conformations are geometrically very similar, with comparable energies differing by only a few kcal mol⁻¹. Figure 2-13.

proper non-crystallographic symmetry (NCS)

Non-crystallographic symmetry that is rotational only and close to a **point group** (closed group) operation. In contrast to the translation periodicity restrictions of crystallographic point groups, also 5-fold and higher NCS rotation axes are allowed. Proper NCS manifests itself in **Patterson self-rotation maps**. Proper crystallographic NCS axes parallel or very near parallel to crystallographic axes that are otherwise obscured in self-rotation maps can manifest themselves in **native Patterson maps**. Figures 11-12, 13-23. Note that proper has a different meaning here than in **proper symmetry operation**.

proper rotation – see plain rotation**proper symmetry operation**

A proper symmetry operation does not change the properties of the **motif** it is acting upon. Only **plain rotations** and **screw axes** are allowed crystallographic symmetry operations for **asymmetric (chiral)** motifs such as protein molecules.

proportional counter

Detector where a photon ionizes a counting gas, generating multiple electrons pulled toward the anode and measured as an electric pulse.

protease (proteinase, proteolytic enzyme)

Enzyme that degrades proteins by hydrolyzing specific peptide bonds between amino acids.

protein

A linear **heteropolymer** built by combination of 20 different common **proteinogenic amino acids**. The term protein is generally used for longer molecules, while shorter stretches (< ~50 residues) not necessarily folding into distinct **tertiary structures** are termed **polypeptide**. Proteins are the major macromolecular constituent of cells.

protein construct

A protein or fragment thereof with a specific sequence, generally engineered and overexpressed, often linked with **affinity tags**, fluorescent probes, and so forth.

Protein Data Bank (PDB)

A curated, world-wide repository of all experimentally determined protein structures, also providing a multitude of online structure analysis tools and database access.

protein engineering

General term for modifying, mostly at the genetic level by **recombinant DNA** techniques, the **sequence** and properties of a protein.

protein stock solution

Aqueous, weakly buffered solution of protein that contains additional components as required to stabilize the protein, plus unknown components acquired and carried through during purification.

proteinogenic amino acids

Proteinogenic amino acids serve as building blocks of **proteins** and are L- α -**amino acids**. There are 20 common proteinogenic amino acids, plus **selenocysteine** and **pyrrolysine**.

proton

Positively charged subatomic particle that forms part of the atomic nucleus. Equivalent to a hydrogen atom stripped of its **electron**, forms in solution a **hydronium ion**.

protonation

Loss of electron and acquisition of a positive charge in aqueous medium.

protozoa

Free-living or parasitic, nonphotosynthetic, single-celled, motile eukaryotic organisms, such as *Paramecium* and *Amoeba*. Free-living protozoa feed on bacteria or other microorganisms.

pseudo-symmetry

Non-crystallographic symmetry with rotation axes parallel to crystallographic axes and/or translations with integer fractions of unit cell vectors that generate higher apparent symmetry and abnormal intensity distributions.

psi backbone torsion angle

The backbone **torsion angle** for the torsion around C α and C of a **residue**. Figure 2-7.

pyrrolysine

A rare **proteinogenic amino acid**, found only in methanogenic **archaea** in **enzymes** that are part of their methane-producing metabolism.

Q**quasi-elastic light scattering – see dynamic light scattering****quaternary structure**

The 3-dimensional arrangement of more than one polypeptide chain (subunits) in a molecular complex.

quaternion

Quaternions form a 4-dimensional normed division algebra over the real numbers. Used as 4×4 matrices in the description of certain 3-dimensional rotations and superpositions.

R**R-anom** (8-18)

The **linear merging R-value** for anomalous pairs of reflections.

R-cryst

Commonly the **linear merging R-value** for reflections from different crystals.

R-Cullis (10-56)

Various linear residuals once used for tracking **heavy atom refinement**.

R-delta (10-4)

Generic difference data merging **R-value**.

R-factor – obsolete see R-value**R-free – see cross-validation R-value** (12-35)**R-free/R ratio** (12-103)**R-int**

Commonly the **linear merging R-value** for identical reflections; based on intensities.

R-Kraut (10-57)

R-value based on linear **closure residual**.

R-meas – see redundancy-independent merging R-value (8-16)**R-merge – see linear merging R-value** (8-15)**R-pim – see precision-indicating merging R-value** (8-17)**R-rim – see redundancy-independent merging R-value** (8-16)**R-sigma** (8-25)**R-sym**

Commonly the **linear merging R-value** for symmetry related reflections.

R-value, R_f -value (7-51)

The normalized **linear residual** between observed and calculated structure factor amplitudes. Expectation value for non-centrosymmetric random structure ~ 0.59 , for centrosymmetric case ~ 0.83 .

R-work (12-35)

R-value for all reflections not belonging to the **cross-validation** data set. *See also* **cross-validation R-value (R-free)**.

racemic mixture, racemate

Mixture of equal parts of **enantiomorphs**.

radiation damage

Resulting from interaction of high-energy **ionizing radiation** such as X-rays, causing bond breakage and formation of **free radicals**. Destroys protein crystals and is routinely counteracted by keeping crystals at **cryogenic** temperature during exposure.

radiation damage-induced phasing (RIP)

An unusual phasing technique that exploits the intensity differences between data recorded from a crystal at different stages of **radiation damage**.

radical

Reactive species featuring unpaired electrons.

radius of convergence – see convergence radius**radius of gyration**

The root mean square distance of the atoms from the **center of mass** of a molecule.

Ramachandran plot – see main chain torsion angle plot**random error**

Deviations from the mean following a **normal distribution**.

random variable

A variable that can take any value in a given interval depending on its probability distribution, independent of past or future observations.

raw moment, k th raw moment

The moments of a distribution about the origin, statistical descriptors used in intensity statistics and defined as $\langle x^k \rangle$ for the k th raw moment. *Compare* **central moment**.

Rayleigh scattering

Elastic scattering electromagnetic radiation by particles much smaller than the **wavelength** of the light, that is, with a **Mie ratio** $x_M \ll 1$.

real space

The 3-dimensional space in which crystallographers and physical objects reside.

real space correlation coefficient (plot), RSCC (7-54)

A scale-factor-independent measure, plotted residue-by-residue, for the agreement between observed (bias minimized) and calculated (model) electron density.

real space molecular replacement

Molecular replacement methods to determine **NCS operators**, applied to electron density maps by **domain rotation functions** and **phased translation functions**.

real space R-value (13-3)

Linear residual between observed and calculated electron density.

real space refinement

Local refinement of the positional parameters of the structure model against experimental or **bias** minimized electron density.

receptor

Any protein inside the cell or at the cell surface that binds a specific signal molecule (ligand) and initiates a response in the cell.

reciprocal lattice

A periodic lattice in **reciprocal space** R^* with basis $[0, \mathbf{a}^*, \mathbf{b}^*, \mathbf{c}^*]$ consisting of **reciprocal lattice points** and reciprocal to real space (crystal lattice) $[0, \mathbf{a}, \mathbf{b}, \mathbf{c}]$. *See also* **reciprocal space**.

reciprocal lattice point

Any point $\mathbf{r}^* = (h\mathbf{a}^* + k\mathbf{b}^* + l\mathbf{c}^*)$ with h, k, l integers in a lattice with basis $[0, \mathbf{a}^*, \mathbf{b}^*, \mathbf{c}^*]$ reciprocal to basis $[0, \mathbf{a}, \mathbf{b}, \mathbf{c}]$. **Reciprocal lattice points** represent sets of parallel equidistant lattice planes with **Miller indices** hkl in a real space lattice or crystal structure.

reciprocal lattice vector

A vector from the origin to a **reciprocal lattice point** (hkl), collinear with the normal vector and interplanar distance vector $\mathbf{d}(hkl)$ of lattice plane hkl .

reciprocal space

A space R^* related to a real space R by the reciprocity conditions of the basis vectors so that the **dyad product** of their **basis vectors** forms the **identity matrix**, that is, in three dimensions $\mathbf{a}^* \cdot \mathbf{a} = \mathbf{b}^* \cdot \mathbf{b} = \mathbf{c}^* \cdot \mathbf{c} = 1$. The diffraction space spanned by the **reciprocal lattice** with basis $[0, \mathbf{a}^*, \mathbf{b}^*, \mathbf{c}^*]$ is a reciprocal to the **real space** with basis $[0, \mathbf{a}, \mathbf{b}, \mathbf{c}]$ and *vice versa*. Real space and reciprocal space share the same origin. Objects or functions are mapped from one space or domain into its reciprocal space or domain by **Fourier transforms**.

reciprocal space interference (G-)function

Function that determines how reciprocal lattice points (structure factors) calculated from different regions of direct space affect each other (11-26).

reciprocal space refinement

Nonlinear refinement procedure in which the minimization target function is based on structure factor amplitudes, which are elements of reciprocal space (i.e. diffraction space). *See also* **restrained reciprocal space refinement**.

recombinant DNA

Any DNA molecule formed by joining DNA segments from different sources.

recombinant DNA methods, technology

Collection of techniques by which DNA segments from different sources are combined to make a new DNA, often called a recombinant DNA. Recombinant DNAs are widely used in the **cloning** of genes, in the genetic modification of organisms, and in production of proteins by overexpression in an **expression host** or **cell-free system**.

rectangular lattice

A 2-dimensional plane lattice with $a \neq b$, $\gamma = 90^\circ$.

redundancy-independent merging R-value (8-16)**refinement**

General term for iterative adjustment of variable **parameters** of a model so that the fit between model and observed data is optimized by minimizing the sum of squared residuals of a refinement target function, generally a **maximum likelihood** residual. Different refinement protocols and **optimization algorithms** exist. *See also* **restrained reciprocal space refinement**.

reflection conditions – see systematic absences**reflections (reflexions) – see X-ray reflection****refractive index**

Measure for how much the propagation speed of a photon is reduced in a medium compared to vacuum. In most materials proportional to square root of the **dielectric constant**, and approaching unity in all materials for **X-ray** radiation.

register error

Model building error in which the main chain conformation is correct, but the sequence assignment and with it the side chains are shifted.

re-indexing

Assignment of new unit cell axes using a basis transformation in the form of a re-indexing matrix.

residual, deviate (7-28)

The difference between two values, often between observed and calculated values or observed and **expectation values**.

residual index – see *R-value* (7-51)

residue

Amino acid **monomer** within a **polypeptide** chain; **side chain** plus peptide **main chain atoms** (Figure 2-19).

resolution (C)

The limit in **diffraction angle** up to which X-rays diffracted by a crystal can be detected. Expressed in smallest interplanar **lattice spacing** d_{min} , sampled, approximately equal (slightly larger than) to the distance at which objects can be distinguished in an electron density map (Sidebar 9-4). The higher the resolution, the smaller the lattice spacing, and the more detailed the **electron density** map. Lattice spacing and diffraction angle are related through the **Bragg equation**. Compare **effective resolution**.

resolution sphere (C)

A sphere (or ellipsoid in the anisotropic case) with radius of $1/d_{min}$ containing the actually observed reflections intersecting the **Ewald sphere**, drawn centered at the reciprocal lattice origin in a Ewald construction.

resonance

Maximal oscillation of a system in response to a stimulus, such as atomic electrons excited by the electric field vector of electromagnetic radiation. A region in the spectrum where absorption and anomalous dispersion effects occur.

restrained reciprocal space refinement

Refinement against structure factor amplitudes where the **sum of residuals squared** includes additional residual terms for the deviations of the model geometry from **stereochemical restraint target values**. See also **restraint weights**.

restraint target values – see **stereochemical restraint target values**

restraint weights

The inverse variance of the respective restraint target distribution.

restraints – see **stereochemical restraints**

restriction nuclease (restriction enzyme)

Enzyme that can cleave a DNA molecule at any site where a specific short sequence of nucleotides occurs. Important for recombinant DNA technology.

rhombohedral

A lattice or unit cell that can be either indexed rhombohedrally primitive with $a=b=c$, $\alpha=\beta=\gamma \neq 90^\circ$, or preferably indexed in **obverse hexagonal** setting as a rhombohedrally centered **Bravais lattice** or cell.

rhombohedral centering

A rhombohedral cell can be described as a triple-sized trigonal cell with rhombohedral *R*-centering vectors $(\frac{2}{3} \frac{1}{3} \frac{1}{3})$ and $(\frac{1}{3} \frac{2}{3} \frac{2}{3})$, lattice type *hR*.

ribosome

Particle composed of ribosomal RNAs and ribosomal proteins that catalyzes protein synthesis using information provided by mRNA (Figure 2-1).

ribozyme

RNA with catalytic activity.

Rice distribution (7-148)

A special case of non-central χ -distribution, the form of the **Sim distribution** and the general, conditional **acentric** structure factor probability distribution.

riding positions

The calculated positions of hydrogen atoms in a molecular structure.

rigid body refinement

Refinement of the position of entire molecules or molecular groups treated as rigid units in a crystal structure. Often used subsequent to **molecular replacement** searches in early steps of macromolecular refinement.

RIP – see **radiation damage-induced phasing**

r.m.s.d., RMSD (7-33)

Root mean square deviation, square root of the **variance** of a distribution.

RMS-Z score (7-37)

Normalized statistical standard score, equals unity when the sample distribution has the same variance as the target distribution, smaller than 1 for tighter distributions, larger for more loose distributions. See also **Z-score**.

RNA (ribonucleic acid)

Polymer formed from covalently linked ribonucleotide monomers. Different functions, such as messenger RNA, ribosomal RNA, transfer RNA. Forms secondary structures that sometimes cause problems in the translation phase during heterologous overexpression.

rotation

Linear transformation or operation by which an object is rotated around a given **rotation axis**. Every rotation can be described by a rotation by one **principal Euler angle** around the **Euler axis**. Expressed as linear matrix equation $\mathbf{x}' = \mathbf{R}\mathbf{x}$ with rotation matrix **R** the **direction cosine matrix**.

rotation axis

Axis around which a general rotation is performed. Compare **crystallographic rotation axis**.

rotation camera – see **rotation method**

rotation function

Rotation functions are used to find the relative orientation of a search molecule in an unknown crystal structure. They can be based on the Patterson rotation function or can be implemented as **maximum likelihood rotation functions**.

rotation matrix

Description of a rotation operation by a 3×3 matrix with special properties. See **direction cosine matrix**.

rotation method

Simplest method of X-ray data collection by exposing a crystal to X-rays while it is rotated in small increments around a single axis. During each rotation increment one **diffraction image** or **frame** is recorded.

rotation search – see **rotation function**

roto-inversion – see **inversion axis**

A combination of crystallographic **rotation axis** with **inversion**. This improper symmetry operation is not allowed in protein crystals.

roto-translation – see **crystallographic screw axis**

S

SAD – see **single-wavelength anomalous diffraction phasing**

salt bridge, ion pair

Charged, non-directional (ionic) interaction between two side chains, for example Glu-Asn.

salting in

Increasing protein solubility by increasing salt concentration, in low-salt regime of protein solubility diagram.

salting out

Exponential decrease of protein solubility due to increasing salt concentration, in high-salt regime of protein solubility diagram.

salvage strategies

Methods applied to modify protein after the fact that crystallization did not succeed, such as **lysine methylation** or enzymatic digests.

sampling probability – see **data likelihood**

satellite reflections

Weak systematic reflections next to Bragg reflections, often indicative of an **incommensurately modulated structure**.

Sayre's equation (10-21, 10-24)

Convolution of **structure factors** in **reciprocal space**, equivalent

to squaring of **electron density in real space**. Used together with **structure invariants** in **direct methods** of phase determination.

SBS standard

The Society for Biomolecular Screening (SBS) provides standards for labware and plastics assuring that their form factors are compatible with SBS standard compliant instruments and robotics.

scalar product (M)

The scalar (or inner, dot) product of two column vectors **a**, **b** is the defined as $\mathbf{a} \cdot \mathbf{b} = ab \cos \gamma$ where γ is the angle enclosed between the vectors and *a* and *b* the **norm** of **a** and **b**.

scale factor (overall, linear)

Brings observed diffraction intensities or structure factor amplitudes onto a common, absolute scale with calculated intensities or structure factor amplitudes. Obtained initially from a **Wilson plot**.

SCALE records

Records in PDB file containing the **deorthogonalization matrix**.

scaling

Procedure of bringing various data sets onto a common relative or absolute scale. *See also* **Wilson scaling**.

scattering diagram

A diagram representing the change of wave vectors (momentum transfer) during a scattering process.

scattering envelope

Commonly used term for the scattering function of an entire (single) molecule.

scattering factor – *see* atomic scattering factor

scattering probability

Probability distribution function describing the spatial distribution of scattered waves or particles. *See also* **atomic scattering factor**.

screw axis – *see* crystallographic screw axis

SDS-PAGE (sodium dodecyl sulfate–polyacrylamide gel electrophoresis)

Type of electrophoresis used to separate proteins by size. The protein mixture to be separated is treated with a harsh negatively charged detergent (SDS) and with a reducing agent (β -mercaptoethanol), before being run through a polyacrylamide gel. The detergent and reducing agent unfold the proteins, free them from association with other molecules, and separate the polypeptide subunits. *See also* **electrophoresis**.

Se-Met labeling

Introduction of Se-Met either by overexpression in methionine synthesis deficient (Met⁻) host cells or by **metabolic inhibition** in Se-Met augmented medium.

second virial coefficient

A measure for the thermodynamic excess energy, measuring pair-wise intermolecular interactions in a solution, available via osmotic pressure or **light scattering methods**.

secondary structure

Sequence of structural elements in a macromolecule with regular local folding patterns, defined in proteins by a specific pattern of **hydrogen bonds** between **main chain atoms** of the protein chain. Most common secondary structure elements are α -**helices**, β -**sheets**, and various **turns** and **loops** connecting them.

seeding (C)

Introduction of external crystallization nuclei in **supersaturated** protein solution, in order to induce **nucleation** and phase separation.

self-adjoint matrix – *see* Hermitian matrix

self-Patterson vectors

Interatomic distance vectors between atoms of the same molecule in a crystal structure. They are typical for a molecule irrespective of the crystal structure it is in, and can therefore be used in Patterson rotation searches for determining the orientation of a molecule.

self-rotation map – *see* Patterson self-rotation map

self-vectors – *see* Patterson self-vectors

selenocysteine

A rare **proteinogenic amino acid**, isomorphous to L- α -**cysteine** but with Se instead of S, found in certain redox enzymes and hydrogenases.

selenomethionine

Isomorphous to L- α -**methionine** but with Se instead of S, introduced by overexpression in medium augmented with Se-Met instead of native methionine. Used as a source of isomorphous **anomalous signal** for **anomalous phasing**. Figure 2-21. *See also* **Se-Met labeling**.

semi-invariants – *see* structure semi-invariants

sequence – *see* primary structure

sequence identity

The percentage of identical residues between two aligned protein sequences.

sequence similarity

The percentage of identical and functionally similar residues between two aligned protein sequences, based on a similarity or substitution matrix.

serial extinctions

Systematic absences that extend along a specific direction in the reciprocal lattice, along a reciprocal axis. The 11 **enantiomorphic pairs of space groups** display the same serial extinctions ($l = 3n$ only for $P3_1$ and $P3_2$, for example) and cannot be distinguished by extinction rules (Table 6-6) alone.

serine

A **polar** L- α -amino acid with a short hydroxyl side chain $-\text{CH}_2\text{OH}$, common target for phosphorylation by kinases.

serine protease

Type of protease that has a reactive serine in the active site.

Shake and Bake (SnB)

A general purpose dual-space direct method program for structure solution, also used for marker atom **substructure** solution.

Shake and wARP

A slow but powerful, iterative **model bias** minimization and model completion protocol based on ARP/wARP combining coordinate perturbation, **omit** techniques, unrestrained maximum likelihood **dummy atom refinement**, and multi-model **map averaging** to generate bias minimized electron density maps. Generally requires data with **resolution** of $\sim 2.5 \text{ \AA}$ or better.

shape complementarity

Fit of molecules to **receptor** based on shape giving rise to **van der Waals interactions** (in addition to specific interactions) which contribute a significant part of the binding energy of drug–receptor interactions

sharpening

Sharpening implies elimination of the attenuation of the structure factor magnitude with increasing scattering angle $\sin \theta/\lambda$ by assumption of point atoms resulting in **normalized structure factor** amplitudes E_h . Maps generated from these E_h as coefficients are sharpened. In practice, somewhat less extreme sharpening (sharpening also leads to increased **truncation ripples**) is desirable, and “mixed” structure factor amplitudes of the form $(E_h^3 \cdot F_h)^{1/2}$ are commonly used.

Sheldrick's rule

States that a resolution of 1.2 \AA or better is required for *ab initio* structure determination of proteins by direct methods.

SHELX programs

Suite of fast and robust programs by George Sheldrick for general and macromolecular crystallography, including experimental phasing, density modification, and refinement.

side chain

Amino acid **monomer** or **residue** sans **main chain atoms**, beginning with C β . Figure 2-19.

side chain torsions

The set of torsion angles in an amino acid side chain.

sigma-A, σ_A

A **maximum likelihood**-derived statistical **weight** (variance) accounting for both *incompleteness* and errors in **structure models**, thus providing more accurate structure factors and thus better Fourier coefficients for **electron density map** reconstruction. Generally estimated from a 2-parameter, scattering angle dependent exponential fit (12-19) using the **cross-validation** data set in order to minimize bias.

sigma cutoff

Truncation of data below a certain signal-to-noise ratio, commonly applied to high resolution shells. Necessary for **difference data**, but not in maximum likelihood refinement.

signal sequence

Short continuous sequence of amino acids that determines the final location of a protein in the cell, for example the N-terminal signal peptides directing nascent secretory and transmembrane proteins to the endoplasmic reticulum.

signal-to-noise ratio (SNR)

The fraction of the measured signal divided by its standard deviation or noise level.

signaling cascade, pathway

Sequence of linked intracellular reactions, typically involving multiple amplification steps in a relay chain, triggered by an activated cell-surface **receptor**.

Sim distribution (7-127)

A probability distribution function accounting for incompleteness in structure models (partial structure), similar in shape to the **Rice function**.

Sim weights (7-127)

Statistical **weights** accounting for incompleteness in structure models (partial structure), expressed in terms of modified Bessel functions of the first kind. Figure 7-16.

similarity transformation

Transformation of type $R_w = AR_c A^{-1}$ between basis systems W and C.

Simplex (Nelder–Mead) algorithm

Non-linear numerical optimization algorithm that does not need derivatives.

simulated annealing

A general technique of energy refinement by successively minimizing the potential energy of a perturbed model that slowly returns to equilibrium following a certain *annealing schedule* $T_1 > T_2 > T_3 \dots T(\text{final})$. The X-ray term as well as the stereochemistry of the model are expressed as potential energy terms. During the cooling the model relaxes into a minimum energy. High **radius of convergence**, particularly when combined with **torsion angle refinement**.

simultaneous equations

A system of n equations with multiple parameters, not necessarily linear.

single anomalous diffraction (SAD) phasing

Phasing technique that utilizes a single **anomalous difference** data set. Needs to resolve two ambiguities: **substructure handedness** and **phase angle ambiguity** in phasing equations, generally by means of **density modification**.

single isomorphous replacement (SIR)

Phasing technique that utilizes a single **isomorphous derivative**. Needs to resolve two ambiguities: **substructure handedness** and **phase angle ambiguity** in phasing equations, generally by means of **density modification** or addition of orthogonal **anomalous signal** (SIRAS).

single isomorphous replacement with anomalous scattering (SIRAS)

Phasing method that uses the additional **anomalous signal** of the

heavy marker atom to break the **phase angle ambiguity** inherent in **single isomorphous replacement** (SIR).

SIR – see single isomorphous replacement**SIRAS – see single isomorphous replacement with anomalous scattering****site-directed mutagenesis**

Technique by which a mutation can be made at a particular site in DNA. Sidebar 4-3.

sitting drop (C)

Crystallization method by **vapor diffusion** where a drop of protein solution plus precipitate sitting on a shelf, post, or depression equilibrates against a well solution of **precipitate** solution. Most common method in high-throughput format. *See also hanging drop*.

skew (skewness)

The third **central moment** of a distribution defining the asymmetry of the distribution.

slope (M)

The slope of a **function** at a given point is given by its first **derivative**. At **extrema** of a function the slope or first derivative is zero. The **curvature** is negative for maxima, positive for minima of a function.

slow cooling protocol

Temperature schedule for simulated annealing optimization in molecular dynamics refinement.

small angle scattering

Scattering experiment on amorphous or liquid sample in which the scattering is limited to forward scattering representing coarse features of the molecular **scattering envelope**.

soaking

Incorporation of a heavy atom, ligand, or other (small) binding partners into an already grown crystal by adding the substance to the mother liquor.

solubility, of proteins

Protein solubility is affected by various reagents such as salts determining ionic strength, small alcohols affecting dielectric constant of the solvent, **poly(ethylene glycol)s** (PEGs) competing for solvent interactions, and net and local charge as affected by pH.

solubility diagram

A diagram of protein concentration versus precipitant concentration depicting the **solubility line** (solubility limit). *See also crystallization diagram*.

solubility line

Line in a phase diagram indicating maximum (protein) solubility and separating the single-phase (protein) solution from a two-phase region (or region of metastability).

solvent (C)

In protein crystallization, generally referring to disordered **mother liquor** in **solvent channels** and intermolecular voids in the crystal structure. Contains all components of the **crystallization cocktail** plus all components present in the **protein stock solution**.

solvent channels

Contiguous channels in protein crystals filled with disordered mother liquor, often along **plain rotation axes** in crystal structures.

solvent content

Percentage of crystals structure occupied by disordered solvent (crystallization **mother liquor**) obtained from **Matthews probabilities**.

solvent flattening

An iterative **solvent modification** method providing phase improvement and **phase extension** based on setting the solvent density to constant value.

solvent flipping

An iterative **solvent modification** method providing phase

- improvement and **phase extension** that is robust against bias from the protein map. Similar in concept to **charge flipping**.
- solvent mask**
Mask of connected real space grid points delineating the protein molecule (density) from the disordered solvent region.
- solvent modification**
General term encompassing methods for phase improvement and phase extension
- solvent polarity**
Solvents consisting of polar molecules such as water have a high polarity; aliphatic molecules such as hexane have a low polarity.
- space group**
Mathematical **group** of symmetry operations, with one to three **generators**, giving rise to the general 230 (65 chiral) different arrangements of motifs in a 3-dimensional periodic crystal structure.
- space group enantiomorph** – see **enantiomorphic pairs of space group**
- space group symbol**
Symbolic representation of the **space group** symmetry and its generators, commonly in the form of the **Hermann–Mauguin symbol** including the **Bravais symbol** and one to three **generators** of the group. The alternative Schönflies notation and Hall notation are less common in macromolecular crystallography.
- spallation neutron source (SNS)**
When a high-energy **proton** is accelerated into a target made of heavy elements such as mercury or uranium, spallation particles including **neutrons** are produced. For every proton striking the nucleus, 20 to 30 neutrons are expelled. The Oak Ridge SNS is presently the most powerful neutron source in the world.
- sparse matrix (C)**
A crystallization experiment with a given combination of reagents and reagent levels, based on prior experience, but not following a rational **experimental design**.
- sparse matrix (M)**
A **sparse matrix** is a matrix populated primarily with zeros. Similar to a scientific meeting.
- sparse matrix refinement**
Minimization algorithm in which only the largest second derivative matrix elements are evaluated.
- special position (C)**
A crystallographic position in a unit cell that is located on a symmetry element or intersection of symmetry elements. A motif on such a position needs to have at minimum the same symmetry as the position where it is located. Asymmetric protein molecules cannot be located on symmetry elements, only on **general positions**. Heavy atoms or solvent molecules can be found on special positions.
- sphere of influence algorithm**
A method of **density modification** that uses a fuzzy cross-over region instead of a sharp solvent mask to define the solvent–protein boundary.
- spherical coordinates**
Spherical coordinates describe a rotation in terms of an azimuthal angle ϕ (horizontal rotation), a lateral angle ψ (up/down rotation), and a rotation κ around the **principal Euler axis** defined by the rotations about ϕ and ψ . Figure 11-5.
- spherical harmonics**
Ubiquitous functions $Y_{l,m}$ that are the angular part of orthogonal solutions to the Laplace differential equation, represented in a system of spherical coordinates.
- spherulites**
Term from mineralogy indicating spherical clusters of needle-like crystals that originate from a **nucleation** core in the center of the spherulite.
- split conformation, side chain**
Two discrete conformations of the same molecule or part thereof, often observed in surface-exposed residues or ligand molecules.
- square lattice (C)**
A 2-dimensional plane lattice with $a = b$, $\gamma = 90^\circ$.
- square matrix (M)**
An $n \times n$ matrix.
- squaring method**
Basis of **direct methods**; **convolution (autocorrelation)** of electron density in direct space is equivalent to squaring of structure factors.
- Src (Src protein family)**
Family of cytoplasmic tyrosine kinases (pronounced “sark”) that associate with the cytoplasmic domains of some enzyme-linked cell-surface receptors (for example, the T-cell antigen receptor) that lack intrinsic **tyrosine kinase** activity. They transmit a signal onwards by phosphorylating the receptor itself and specific intracellular signaling proteins on tyrosines.
- S-SAD** – see **sulfur single anomalous diffraction phasing**
- standard deviation (standard uncertainty)**
Square root of **variance**, measure for width of a probability distribution.
- static light scattering, MALS**
Multi-angle light scattering (MALS) method for molecular weight determination. Also provides the second thermodynamic virial coefficient as a measure for intermolecular interaction.
- statistical weight**
A statistical measure for the reliability of observations or computed parameters. Examples are **variance**, **sigma-A**, and **restraint weights**.
- steepest descent minimization**
Iterative multivariate minimization algorithm that uses only the first derivatives or gradients to determine the downward path.
- stereochemical restraint target values**
Expectation values and **variance** of stereochemical parameters such as bond lengths and angles, torsion angles, planarity of groups and rings, and chiral volumes.
- stereochemical restraints**
Provide a means of keeping the model stereochemistry in physically reasonable bounds in macromolecular refinement because the experimental **data-to-parameter ratio** is generally too low for unrestrained refinement. Most restraints are implemented as sum of residuals squared reflecting model deviations from stereochemical restraint target values.
- stereographic projection**
A representation of a sphere on a plane, using a stereographic net (Figure 11-12). Historically used for experimental determination of **crystal class** and frequently in the representation of a rotation, particularly **Patterson self-rotations**, in **spherical coordinates** (lateral angle ψ , azimuthal angle ϕ) in sections with different κ -angles (**principal Euler angles**).
- stick model**
A model of a molecular structure in which the covalent bonds are represented as sticks, with the atoms located at the intersection of the bond sticks. *Compare ball and stick model.*
- stochastic search (optimization) methods**
Search or optimization algorithms that use random starting parameters combined with efficient algorithms such as evolutionary programming or Metropolis-Monte Carlo to improve the parameters.
- streak seeding**
Introduction of **nucleation** seeds by swishing a fiber with **microseeds** through a protein crystallization drop.
- structure amplitude** – see **structure factor amplitude**
- structure factor**
A complex function $F(hkl)$ (interpreted as a transversal electromagnetic wave) whose magnitude $F(hkl)$ is proportional to the square root of the observed **intensity** of a diffraction spot or reflection with **reciprocal lattice** indices hkl . $F(hkl)$ is the sum of

all atomic scattering contributions in a **unit cell** in the direction defined by hkl .

structure factor amplitude

The norm, magnitude, or amplitude of the complex structure factor **F**, corresponding to the **amplitude** to the diffracted electromagnetic X-ray wave.

structure factor distribution

Probability distribution function showing the probability for structure factors (or structure factor amplitudes) to have a certain value. Can be unconditional (centric and acentric **Wilson distribution**), or conditional on a model (centric **Woolfson**, acentric **Sim**, or **Rice distribution**).

structure guided drug design

A method of rational drug design where receptor–drug interactions are improved and designed based on knowledge of the receptor, ligand, or a receptor–ligand complex structure.

structure invariants

Structure invariants are **triplet relations** independent of the origin of the structure. Used in **direct methods** phasing. *Compare structure semi-invariants.*

structure model – *see* **model** (C)

structure semi-invariants

Structure semi-invariants are **triplet relations** invariant in relation to positions in the cell with the same point group (i.e. the structure factors have same amplitude but different phase). Used in **direct methods** phasing.

substrate analog

Molecule similar to a natural substrate, often a non-hydrolyzable or non-processable form of the substrate.

substructure

A partial assembly of atoms of a structure, generally referring to the **marker atom** substructure in **experimental phasing**.

substructure handedness

Due to symmetry of the **Patterson space**, the **substructure handedness** is generally undefined. When the substructure handedness is changed during phasing, the **screw axis** of an **enantiomorphic pairs of space group** must also be changed to the other enantiomorph of the pair. There are three exceptions to the **inversion** about the origin: In **chiral space groups** $I4_1$, $I4_122$, and $F4_32$, the origin is not located on the enantiomorph axis, and the center of inversion does not coincide with the origin. Their inversion operators are $(-x, \frac{1}{2}-y, -z)$, $(-x, \frac{1}{2}-y, \frac{1}{4}-z)$, $(\frac{1}{4}-x, \frac{1}{4}-y, \frac{1}{4}-z)$, respectively.

suicide substrate

A substrate, ligand, or drug that binds covalently to its **receptor**.

sulfhydryl (thiol)

Reactive chemical group $-SH$ containing sulfur and hydrogen; present in the amino acid cysteine and other molecules. Two sulfhydryls can join to produce a **disulfide bond**. Often covalently modified and subject to **radiation damage**.

sulfur single anomalous diffraction (S-SAD) phasing

Phasing technique that utilizes a single **anomalous difference** data set resulting from weak anomalous signal of native sulfur atoms. Needs to resolve two ambiguities: **substructure handedness** and **phase angle ambiguity** in phasing equations, generally by means of **density modification**.

sum of residuals squared (SRS) (7-34)

Quantity minimized during optimization procedures.

superposition

Minimization of RMSD between corresponding atoms of similar structures. Problem is the detection of the correspondence at low similarity.

supersaturation

Thermodynamically **metastable** state reached by increasing the protein concentration beyond its equilibrium solubility limit.

superstructure – *see* **commensurately modulated structure**

surface entropy reduction

A method that potentially increases probability of crystal formation by reducing the adverse surface entropy term by replacing high-entropy residues such as lysine with smaller and more rigid residues.

symbolic operators (C)

Representation of a symmetry operation not through combined rotation and translation matrix but by symbols representing the coordinates of the object after application of the respective operation. *See also* **matrix operators**.

symmetric matrix

Matrix that is identical with its **transpose**.

symmetry

Arrangement of like objects in a defined fashion.

symmetry averaging – *see* **density averaging**

symmetry axis

The axis around which a **rotation** or **roto-translation** (screw operation) occurs.

symmetry center – *see* **inversion** (C)

symmetry element – *see* **symmetry operation**

symmetry equivalents

Identical symmetry related points (equipoints), objects, or reflections.

symmetry operation

An affine **transformation** or operation acting on an object or **motif** that generates identical copies of the object, that is, the physical properties of the object must remain invariant with respect to any symmetry operation. Represented as a matrix equation $\mathbf{x}' = \mathbf{R}\mathbf{x} + \mathbf{T}$.

symmetry-related molecules

Identical copies of a given molecule that are related by **symmetry operations**.

symmetry-related reflections

Reflections that are symmetry related in reciprocal space as a result of crystal symmetry. Symmetry related reflections have invariably identical reflections and defined phase relations if the symmetry operation includes translational elements, $\varphi_{hR} = \varphi_h - 2\pi\mathbf{hT}$.

synchrotron (storage ring, source)

A machine that keeps electrons circling at relativistic speeds, providing a source of tunable high energy **electromagnetic radiation (synchrotron radiation)** ranging from far UV to hard X-rays (Figure 8-5).

synchrotron radiation

High energy electromagnetic radiation emitted from a synchrotron storage ring by **bending magnets** and **insertion devices** such as **wigglers** and **undulators**. The advantage of synchrotron X-ray sources is that from wide bandwidth of the emitted **brilliant** and polychromatic radiation, narrow **X-ray photon** energies can be selected by suitable X-ray optics and **monochromators**. Tunable X-rays are necessary for **multi-wavelength anomalous diffraction (MAD)** experiments and for optimal selection of wavelength in **anomalous diffraction** experiments in general.

synergistic

Term used to describe a situation where different entities cooperate or interact toward a final outcome.

systematic absences

Reflections that have invariably zero intensity due to the presence of translational symmetry elements (in **screw axes** and **Bravais translations**). Screw axes cause **serial extinctions** and Bravais translations cause **zonal** or **integral extinctions**.

systematic error

Non-random error in measurement affecting a deviation of the **mean** or **expectation value** from the unknown true value. *See also* **accuracy**.

T

tangent formula (10-28)

Relates phase angles between sets of certain **normalized structure factors**. It is used for **phase extension** and phase refinement in **direct methods**.

target values – see **stereochemical restraint target values**

Taylor series (7-70)

Expansion of a function in a series of derivative terms.

tedium (Latin *taedium*: disgust, boredom)

Occurring during late stages of macromolecular **model building** and **refinement**, ultimately terminating macromolecular refinement.

temperature factor – see **Debye–Waller factor**

template convolution

Template matching technique for finding small parts of an image or electron density which match a template image or density.

tensor

Usually a second rank tensor or 3×3 matrix describing an anisotropic physical property, obtained as the dyad product of an anisotropic property vector, for example anisotropic displacement.

tertiary structure

Complex 3-dimensional structure of a folded polymer chain, especially a protein or RNA molecule.

tetrad (C)

Name for 4-fold **rotation axis**.

tetragonal

Lattice or crystal structure with cell parameters $a = b \neq c$, $\alpha = \beta = \gamma = 90^\circ$, with primitive (*tP*) or *I*-centered (*tI*) lattice type and 4-fold internal minimal symmetry.

TEV protease

Very specific and relatively efficient protease of the tobacco etch virus, used commonly in His-tagged variants for tag removal.

thermal diffuse scattering (TDS)

Thermal diffuse scattering is the intensity that is scattered by crystals outside of discrete Bragg reflections due to deviations from the perfect order such as thermal motion and various static or dynamic defects of **atomic displacements**. Diffuse scattering features depend on the correlation functions of the deviations.

thermal ellipsoid – see **anisotropic displacement ellipsoid**

thiol – see **sulfhydryl**

Thomson scattering

Thomson scattering is the elastic scattering of **electromagnetic radiation** by a charged particle such as an **electron**.

three-ten helix (3_{10} helix)

A rare, tightly packed helical secondary structure element with backbone hydrogen bonds from residue n to $n + 3$.

threonine

A **polar** L- α -amino acid with asymmetrically branched hydroxyl side chain $-\text{CH}(\text{OH})\text{CH}_3$, absolute configuration (2S, 3R), common target for phosphorylation by kinases.

time-resolved X-ray diffraction

Rapid collection of data of a crystal or molecules triggered to assume a specific conformational state. Used in exploration of enzyme or electron transport kinetics.

TLS refinement

Macromolecular **refinement** protocol in which the displacement of whole groups of atoms or parts of the molecule such as domains are parameterized as a combination of **translation**, **libration**, and **screw** motion. Improved representation of molecular flexibility and mobility, requiring only 20 parameters per TLS group, therefore usable also at much lower resolution than individual **anisotropic B-factor** refinement.

topoisomerase (DNA topoisomerase)

Enzyme that binds to DNA and reversibly breaks a phosphodi-

ester bond in one or both strands. Topoisomerase I creates transient single-strand breaks, allowing the double helix to swivel and relieving superhelical tension. Topoisomerase II creates transient double-strand breaks, allowing one double helix to pass through another and thus resolving tangles.

torsion angle

The relatively weakly restrained side-chain and main-chain dihedral angles are called torsion angles as compared with general **dihedral angles**.

torsion angle refinement (parameterization)

Refinement protocol in which instead of **atomic coordinates** the **torsion angles** are refined. Requires fewer parameters and has a large **radius of convergence**. Used for refinement of **low resolution structures** and in initial stages of refinement of **molecular replacement** models.

trace (M)

Sum of diagonal elements of a matrix.

trans

On the other or opposite side. *Compare cis*.

transcription (DNA transcription)

Copying of one strand of DNA into a complementary RNA sequence by the enzyme RNA polymerase.

transfection

Introduction of a foreign DNA molecule into a host cell. Usually followed by expression of one or more genes in the newly introduced DNA.

transform – see **Fourier transform**

transformation (B)

Insertion of new DNA (e.g. a plasmid) into a cell or organism, such as into competent *E. coli*.

transformation, affine (M,C)

Function or mapping in vector space that can be expressed as a **linear transformation** plus a translation. A symmetry operation containing a rotation plus a translation is an affine transformation.

transformation, linear (M,C)

Function or linear mapping in vector space that can be expressed as a matrix and preserves vector addition and scalar multiplication. Scaling and rotations are linear transformations.

translation (C)

Linear **transformation**, mapping an object in the same orientation onto a different position, $\mathbf{x}' = \mathbf{x} + \mathbf{T}$.

translation (RNA translation)

Process by which the sequence of **nucleotides** in a **messenger RNA** (mRNA) molecule directs the combination of amino acids into a protein. Takes place at and is catalyzed by **ribosomes**.

translation function

Function used to determine the location of correctly oriented search probe molecule in the unit cell. The search must cover the **Cheshire cell**, and can be a Patterson-based fast translation function or a maximum likelihood translation function.

translation-libration-screw parameterization – see **TLS refinement**

translational non-crystallographic symmetry (NCS)

Non-crystallographic symmetry that is solely or largely translational does not comply with closed **group** limitations. Translational NCS parallel or near parallel and in integer fractional relation to unit cell vectors manifests itself as **commensurately modulated superstructures** and in **native Patterson maps**.

transmembrane helix

A hydrophobic helix of a membrane protein that extends across the cell membrane in contact with the hydrophobic tails of the phospholipid molecules.

transmembrane protein

Integral membrane protein that extends through the lipid bilayer,

generally with functional part present on both sides of the membrane.

transpose (M)

A matrix **M** has a transpose **M^T** with diagonally swapped elements.

triad (C)

Name for 3-fold **rotation axis**.

triclinic

Lattice or crystal structure in which all cell parameters are different, **lattice type** *aP*, *agonic*, with no internal minimum symmetry.

trigonal

Lattice or crystal structure with cell parameters $a = b, c, \alpha = \beta = 90^\circ, \gamma = 120^\circ$, with primitive (*hP*) or *R*-centered (*hR*) lattice type and 3-fold internal minimal symmetry.

triple vector product

Defines the volume *V* of a parallelepiped (or **unit cell**) spanned by vectors **a, b, c** in a right-handed system as positive: $V = \mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})$.

triplet relations (10-26)

Fundamental to direct methods, stating that if all reflections in the reflection triplet $F_{\mathbf{h}}, F_{\mathbf{k}},$ and $F_{\mathbf{h-k}}$ are very large, then the phase relation $\varphi_{-\mathbf{h}} + \varphi_{\mathbf{k}} + \varphi_{\mathbf{h-k}} \approx 0$ is fulfilled.

truncation ripples – see Fourier truncation ripples**tryptophan**

Largest, aromatic and polar L- α -amino acid with indole ring in side chain $-\text{CH}_2(\text{C}_8\text{NH}_6)$; can act as hydrogen bond donor and participate in π -stacking interactions.

turn

Short secondary structure elements with a distinct hydrogen bonding pattern that facilitate tight changes of direction in the polypeptide chain. *Compare loop.*

twin operator

Symmetry operator that relates orientation of twinned domains in a twinned crystal.

tyrosine

Aromatic, polar L- α -amino acid with side chain $-\text{CH}_2(\text{C}_6\text{H}_4)\text{OH}$ that structurally participates in hydrogen bonds and π -stacking interactions.

tyrosine kinase

Enzyme that phosphorylates the **tyrosine** residues of certain proteins.

U**undulator**

Insertion device consisting of a periodic structure of magnets alternating along its length. Electrons traversing the periodic magnet structure are forced to undergo oscillations and therefore radiate. The X-rays produced in an undulator are very intense with narrow energy **bandwidth** and are collimated in the orbit plane of the synchrotron ring.

uniaxial system – see unique axis**unimodal**

A function or probability distribution with a single peak or extremum.

unique axis (C)

Special single axis parallel to minimal symmetry element, **b** in monoclinic, **c** in trigonal, tetragonal, and hexagonal systems.

unique reflections

The reduced set of recorded reflections belonging to one and the same asymmetric unit of reciprocal space.

unit cell

The transitionally repeating building blocks of a crystal, comprising a **unit lattice** filled with motifs according to the **space group** symmetry of the crystal.

unit cell parameters

The dimensions (norm, magnitude) *a, b, c*, of the **unit cell** vectors

a, b, c, and the enclosed angles $\alpha, \beta,$ and γ .

unit cell vector

The three vectors **a, b, c**, enclosing angles $\alpha, \beta,$ and γ that form the **basis** of the **crystal system**.

unit cell volume – see triple vector product**unit lattice**

Mathematical construct spanned by three lattice vectors **a, b, c**, enclosing angles $\alpha, \beta,$ and γ that form the **basis** of the lattice.

unit vector

A **vector** in a vector space whose length (**norm**) is the unit length, 1.

unitary structure factor (7-108)

Structure factor amplitudes $0 \leq U(\mathbf{h}) \leq 1$ reflecting the scattering of point atoms, that is, adjusted for **B-factor** attenuation and **atomic scattering factor** attenuation, but not **epsilon-factor** corrected. *Compare normalized structure factors.*

univariant

Function or probability distribution depending on a single variable.

unrestrained refinement

Refinement that uses diffraction data only and does not incorporate knowledge-based restraints. Rarely possible in macromolecular refinement, requires atomic resolution data.

usage bias

In the absence of a count for negative results, success rates (e.g. percentage successful trials) cannot be properly quantified, and anything that is used more frequently will likely deliver more positive results regardless of its actual success rate.

V**valine**

Hydrophobic L- α -amino acid with a $-\text{CH}(\text{CH}_3)_2$ side chain.

van der Waals (vdW) interaction

Interatomic interaction with potential curve following the **Lennard-Jones potential** (Figure 12-18). Strong and rapidly increasing repulsive component once atoms come closer than the van der Waals contact distance, but only weakly attractive force for distances larger than the van der Waals contact distance. Van der Waals repulsions determine the **main chain torsion angle** distribution, giving rise to typical secondary structures such as helices and sheets.

vapor diffusion

A method of crystallization in a closed system where the protein solubility is decreased by exchange of water vapor from a drop of protein plus precipitant into a **hygroscopic** well solution. Used in **hanging drop** and **sitting drop** setup.

variance (7-29)

Fundamental and universal statistical measure for the width of a **probability distribution** $\sigma_n^2 = \langle (n - \langle n \rangle)^2 \rangle = \langle n^2 \rangle - \langle n \rangle^2$.

variance, of sample distribution (7-32)**vector (B) – see expression vector****vector (C)**

Entity endowed with both length and direction, represented as a 1-dimensional matrix. Can be a row or column vector.

vector product (M)

The vector product or cross-product $\mathbf{a} \times \mathbf{b}$ of two column vectors **a, b** defines a third vector **c** perpendicular to the plane spanned by **a** and **b** with its **norm** defined by $c = |\mathbf{c}| = |\mathbf{a} \times \mathbf{b}| = ab \sin \gamma$.

virtual ligand screening

In-silico method of screening small molecules by means of **ligand docking**.

virus

Particle consisting of nucleic acid (RNA, retrovirus, or DNA) enclosed in a protein **capsid** and capable of replicating within a host cell and spreading from cell to cell.

vitreous

Amorphous, non-crystalline, glass-like.

W

Watson–Crick base pair – *see* **base pair**

wave vector

Propagation vector of **electromagnetic wave**, vector (cross) product of perpendicular electric field vector and magnetic field vector, with magnitude inverse to the **wavelength** or directly proportional to the **energy** of the wave.

wavelength

The extent of a transversal wave between two points of maximum amplitude (or any repeating point). Inversely related to energy or frequency. *See* **wavelength–energy conversion**.

wavelength–energy conversion

Wavelength in Å = 12397.639/Energy in electronvolts (eV).

wedge

A sector of reciprocal space.

weight (C) – *see* **statistical weight**

weighted average (7-55)

Variance-corrected average (mean) value.

Weiss indices – *see* **lattice planes**

Weissenberg camera

An ingenious mechanical device of early crystallography to record the diffraction spots on circular film so that reciprocal layer lines are separated on typical curves. Still used in some detectors in conjunction with special cylindrical image plates.

white line

An additional peak on and at the **X-ray absorption edge**, resulting from electronic transitions into unoccupied higher states.

white radiation

Polychromatic radiation, radiation of a broad energy **bandwidth**, such as **Bremsstrahlung** or synchrotron radiation from a **bending magnet** or **wiggler**.

wiggler

Insertion device consisting of a periodic series of magnets in a Halbach array design to laterally deflect or wiggle the electron beam inside a **synchrotron storage ring**, thereby generating **brilliant** X-ray radiation of a broad energy **bandwidth**.

Wilson distribution

The unconditional probability distribution for structure factor amplitudes or intensities, different for centric and acentric reflections.

Wilson-like maximum likelihood rotation function (11-54)

Basic likelihood rotation function based on the unconditional Wilson structure factor distribution.

Wilson plot (7-121)

The logarithm of mean observed intensity over mean absolute intensity, binned in resolution shells and plotted against a function of $1/d^2$. The intercept of a regression line through the linear high resolution part of the graph provides the linear **scale factor** k as the intercept and the **isotropic overall B-factor** from the slope. Figure 7-15.

Wilson scaling

Used to bring diffraction data onto a common absolute scale, based on the total scattering estimated for the unit cell from the atomic scattering factor contributions. Requires a **scale factor** and an initial **overall B-factor** accounting for attenuation of scattering by **atomic displacement**, obtained from a **Wilson plot**.

Wolfson distribution (7-148)

The conditional structure factor probability distribution for **centric** reflections.

world coordinates

The **Cartesian coordinates** of **real space** harboring crystallographers and their objects.

X**X-ray(s)**

Term coined by Wilhelm Conrad Röntgen for ionizing radiation in the keV energy range (Ångstrom wavelength range).

X-ray absorption edge

Above certain element-specific energies X-ray absorption rapidly increases, forming an edge in the absorption spectrum. Absorption edges have a fine structure, and in their vicinity atomic scattering factors contain significant, wavelength-dependent **anomalous scattering** contributions.

X-ray absorption near edge structure

Region of the X-ray absorption spectrum in the close vicinity of an absorption edge. Typical features are **white lines**.

X-ray crystallography

Fundamental technique for the exploration of the atomic and molecular structure of matter. Based on discrete **diffraction** of X-rays by periodically ordered assembly of atoms or molecules, that is, a single **crystal** or crystalline matter in general.

X-ray diffraction

Process of interaction of **X-rays** (high energy **electromagnetic radiation**) with the **electrons** of periodically arranged atoms and molecules in **crystals**.

X-ray emission line – *see* **X-ray fluorescence**

X-ray excitation scan

An absorption edge scan obtained by measuring the intensity of the fluorescence excitation (instead of the **X-ray absorption** directly).

X-ray fluorescence

Emitted from materials when an electron from a higher shell fills a hole (generated by high energy X-ray or electron bombardment) in a tightly bound core electron shell. In **X-ray generators**, typically the strong K-shell fluorescence emission lines are used. Figure 8-2.

X-ray generator

Instrument that generates X-rays by bombarding water-cooled anode material (such as Cr, Cu, or Mo) with high energy electrons of ~40–100 keV. Generates continuous **Bremsstrahlung** as a result of electron deceleration, superimposed with about 100–1000 times as intense **characteristic X-ray fluorescence** emission lines with energies specific for the anode material.

X-ray optics

General term for devices to affect properties and geometry of X-rays, such as monochromators, filters, X-ray mirrors, and collimators.

X-ray reflection

Diffraction spot recorded on a film or detector resulting from diffracted X-rays, interpreted as reflections of X-rays on crystal **lattice planes** with **interplanar spacing** $d(hkl)$ related to diffraction (reflection) angle by the **Bragg equation**.

X-ray scattering

Fundamental process of interaction of **electrons** with **X-rays**. Term also used for describing experimental techniques that do not require crystalline matter. *Compare* **X-ray diffraction**.

X-ray sources

X-ray radiation generating devices, generally by (i) acceleration of electrons (**synchrotron** radiation and **Bremsstrahlung**), (ii) characteristic **X-ray fluorescence** (laboratory **X-ray generators**), (iii) radioactive decay (flood field sources), or (iv) inverse Compton effect.

X-ray term

The residual in a refinement that depends on the experimental diffraction data only.

X-ray weights (12-49)

Weight given the X-ray residual relative to the geometry restraint residual in restrained reciprocal space refinement.

XANES – *see* **X-ray absorption near edge structure****xenobiotic**

A chemical which is found in an organism but which is not normally produced by it or expected to be present in it, for example drug molecules. *See also* **cytochrome P-450**.

XPLOR

A versatile high level programming language and program package for macromolecular refinement, superseded by **CNS** (Crystallography and NMR System) by Axel Brunger and colleagues. Unique feature simulated annealing torsion angle refinement.

Y**yeast**

Common name for several families of unicellular fungi. Among the simplest of eukaryotes, used as efficient expression hosts for eukaryotic proteins. Generally capable of most posttranslational modifications, secretion, and disulfide formation.

Yeates–Padilla plots

Cumulative probability plot of pair-wise local intensity relations of reflections, revealing (also perfect) **hemihedral twinning**.

Z**Z-score**

Normalized residual, statistical standard score measuring the deviation of values from their **expectation value** in units of **standard deviations** of their sampling distribution. *See also* **RMS-Z-score**.

zinc finger

DNA-binding structural motif present in many gene regulatory proteins. All zinc finger motifs incorporate one or more Zn^{2+} ions. Figure 2-46.

zonal absences, extinctions

Systematic absences caused by translations in certain planes; for example, *C*-centering causes zonal absences for reflections $hk \neq 2n$ (odd).

zone (C)

Set of planes in a lattice whose *intersections* are parallel; often used as a general term for a set of parallel planes.

zone axis

Direction of the intersection of a **zone** of lattice planes hkl , indicated by square brackets $[hkl]$. Example: all lattice planes $hk0$ have a zone axis $[001]$, and their reflections lie in the $hk0$ reciprocal lattice plane.