Human Insulin Genome Sequence Map, Biochemical Structure of Insulin for Recombinant DNA Insulin

Chiranjib Chakraborty* and Ashish A. Mungantiwar

Departmet of Biotechnology and Medical science, Macleods Pharmaceuticals Ltd, Andheri(East), Mumbai-400059, India

Abstract: Insulin is a essential molecule for type I diabetes that is marketed by very few companies. It is the first molecule, which was made by recombinant technology; but the commercialization process is very difficult. Knowledge about biochemical structure of insulin and human insulin genome sequence map is pivotal to large scale manufacturing of recombinant DNA Insulin. This paper reviews human insulin genome sequence map, the amino acid sequence of porcine insulin, crystal structure of porcine insulin, insulin monomer, aggregation surfaces of insulin, conformational variation in the insulin monomer, insulin X-ray structures for recombinant DNA technology in the synthesis of human insulin in *Escherichia coli.*

INTRODUCTION

Since the early 1980s, a radical change in medical society has occurred through the use of recombinant DNA technology to produce pharmaceutical products like the discovery of penicillin to produce antibiotics. Several diseases can be treated with the recombinant products, for instance, diabetes (lack of insulin), dwarfism (lack of human growth hormone or HGH), and hemophilia A (lack of blood-clotting factor VIII). Supplements of recombinant proteins (e.g., insulin, HGH, blood-clotting factor VIII) can be used in the aforementioned diseases, respectively. Now some of the human proteins that have been synthesized from genes cloned in bacteria and/or eukaryotic cells are being used as remedies for the human diseases [1]. For example, in 1977 the first synthesis of a functional polypeptide product from recombinant DNA technology was the production of hormone somatostatin in *Escherichia coli* using pBR322 as a cloning vector [2]. In 1979, human growth hormone was directly expressed in *Escherichia coli* using pBR322 as a cloning vector [3]. In 1981, a human gene encoding for interferon could be expressed in yeast [4]. Nevertheless, the first commercial application was the microbial production of human insulin [5,6].

Normally, insulin is synthesized from beta-cells of the Islet of Langerhans in the pancreas and controls the level of glucose in the blood [1]. The lack of insulin due to defective insulin genes may lead to a complex of symptoms called diabetes mellitus which can be lethal if not treated. In a normal person, insulin is produced upon induction by high blood sugar and released into the blood stream [6]. Insulin will act as a signaling molecule by binding to the insulin receptors on the surface of a cell to generate the signals for movement of glucose transporters to the cell membrane. These glucose transporters will aggregate into helical structure creating channels for entrance of glucose molecules into the cells. In the absence of insulin, the body cannot take up blood glucose, and glucose is abruptly accumulated at high level in the blood. Although supplementing the patient with insulin from other sources such as from pig insulin and cow insulin can treat this genetic disease, the problems do occur [1]. One problem is that as animal insulin and human insulin are slightly different, allergic reaction may occur in some patients. Another problem is that purification of animal insulin is difficult and microbial contamination is hardly avoidable. However, the recombinant DNA technology has led to a new choice for production of human insulin in large quantity by introducing human insulin genes into the bacteria and synthesizing human insulin in that bacteria. This achievement has been reported since the early 1980s by research group led by Goeddel and his colleagues [3,5,7].

BACKGROUND INFORMATION

The human insulin can be produced in bacteria by recombinant DNA technology because of its two advantageous characteristics [3,5,7]. First, human insulin does not require the post-translational modification such as glycosylation (addition of sugar), so the recombinant insulin synthesized by a bacterium should be active. The second advantage concerns the size of the molecule. Insulin is a small protein comprising of two polypeptide chains (**A-chain** of 21 amino acids and **B-chain** of 30 amino acids) linked together by a **disulfide bonds** between the cysteine amino acids [1,6]. Before becoming an active mature insulin, the insulin is produced from pancreatic cells as preproinsulin, which contain four segments; (1) an N-terminal signal sequence of 16 amino acids, (2) a B-chain of 30 amino acids, (3) a C-chain of 33 amino acids, and (4) an A-chain of 21 amino acids [6]. During the post-translational modification, the N-terminal and the C-peptide are cleaved

*Address correspondence to this author at the Departmet of Biotechnology and Medical Science, Macleods Pharmaceuticals Ltd ,3rd floor,Atlanta Arcade, Moral Church Road, Andheri(East), Mumbai-400059 India ; Tel: +91-22-8213339/8214636/8214637/8214638; Fax: +91-22-821-65 99; E-mail: drchiranjib@yahoo.com

1389-5575/03 $41.00+.00

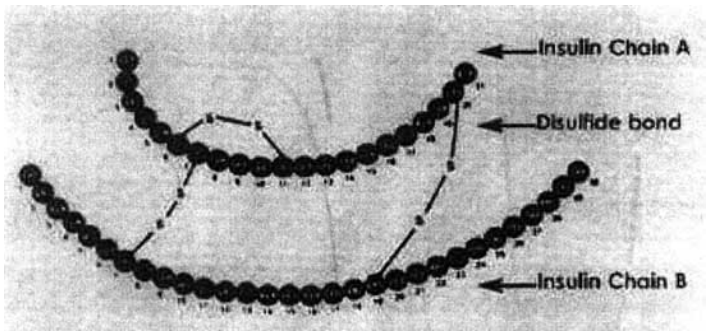


Fig. (1). Insulin-A chain and B chain and disulfide bond.

off, and the disulfide bonds are formed between the A-chain and the B-chain.

HUMAN INSULIN GENOME SEQUENCE MAP

The genetic code for insulin is found in the DNA at the top of the short arm of the eleventh chromosome. It contains 153 nitrogen bases (63 in the A chain and 90 in the B chain).

Currently, molecular genetics and genetic engineering are widely open to many laboratories around the world. The human insulin gene (4044 bp) consists of three exons and two introns coding for a signal peptide, a B-chain, a C-chain, and an A-chain. Present evidence favors a single insulin gene per haploid genome; however, allelic and polymorphic variation are conspicuous [8,9,10]. More information about human insulin nucleotide sequence has been reviewed by some other scientists briefly [11, 12, 13]. For human (synthetic) insulin gene, the nucleotide sequence is only 351 bp [7,14,15]. The TIGER Human Gene Index (HGI) provides insulin gene with sequence (THC479288, THC479502,HT3739,NP078669 etc.) and putative ID which is showed in table 1.

Table 1. The TIGER Human Gene Index (HGI) for Insulin Gene with Sequence and Putative ID

Sequence	Putative ID
THC479288	MAC25; prostacyclin-stimulating factor, PGI2-stimulating factor, PSF [human, cultured diploid fibroblast cells, Peptide, 282 aa]; insulin-like growth factor binding protein 7; PGI2-stimulating factor; PSF
THC479502	IGF-BP 4; insulin-like growth factor binding protein 4; insulin-like growth factor binding protein 4 (GB:U20983); insulin-like growth factor binding protein-4
THC479738	BAP2-alpha protein; insulin receptor substrate protein of 53 kDa (a longer form); Fas-ligand associated factor 3; BAP2-beta protein; insulin receptor substrate protein of 53 kDa (a shorter form); BAI1-associated protein 2
THC480293	putative insulin-like growth factor ii associated protein {Homo sapiens} SP|P09565|IG2R_HUMAN PUTATI
THC480316	precursor polypeptide (AA -24 t0 82) (423 is 1st base in codon); insulin-like growth factor II; insulin-like growth factor II (78 AA); put. IGF-II precursor (175 is 1st base in codon); insulin-like growth factor II, IGF-II [human, small cell lung cancer c
THC480319	IGF-II precursor; insulin-like growth factor II precursor; preproinsulin-like growth factor II, domains A-E
THC481007	insulin-like growth factor binding protein 5; [Human insulin-like growth factor binding protein 5 (IGFBP5) gene], gene product
THC482768	insulin-like growth factor I
THC482769	insulin-like growth factor IB; IGF-1b; insulin-like growth factor IB prepropeptide; pot. IGF-I precursor; insulin-like growth factor-I; put. IGF-I precursor (169 is 1st base in codon)
THC482770	insulin-like growth factor I; IGF-1a; insulin-like growth factor I precursor; IGF-I; insulin-like growth factor 1A precursor; insulin-like growth factor precursor; insulin-like growth factor precursor IA
THC482836	oxytocinase splice variant 2; placental leucine aminopeptidase; leucyl/cystinyl aminopeptidase; oxytocinase/insulin-responsive aminopeptidase, putative variant 2; leucine aminopeptidase, placental; oxytocinase splice variant 1; oxytocinase/insulin-respons
THC484008	glucose insulinotropic peptide receptor; IAPP precursor; IAPP precursor (AA 1-89); GIP receptor; islet amyloid polypeptide; GIPR_HUMAN; glucose-dependent insulinotropic polypeptide receptor; GIP receptor [Homo sapiens]; islet amyloid polypeptide (IAAP); i
THC484787	ribosomal protein S6 kinase 3; insulin-stimulated protein kinase 1; protein kinase 1, insulin-stimulated; ribosomal protein S6 kinase, 90kD, polypeptide 3
THC485756	glucose transporter 4; insulin-responsive glucose transporter; solute carrier family 2 (facilitated glucose transporter), member 4

(Table 1) contd.....

Sequence	Putative ID
THC490180	adapter protein; similarto mouse growth factor receptor-binding protein Grb10.; insulin receptor binding protein; growth factor receptor-bound protein 10 long form; insulin receptor binding protein GRB10
THC490181	adapter protein; SH2 domain containing protein; insulin receptor inhibitor, muscle; growth factor receptor-bound protein 10 short form; alternatively spliced variant of insulin receptor binding protein GRB10
THC490565	insulin-like placentin; EPIL; insulin-like 4 precursor; early placenta insulin-like peptide (EPIL); placentin
THC492627	insulin-like growth factor binding protein 3; insulin-like growth factor binding protein 3; insulin-like growth factor-binding protein; growth factor-binding protein-3
THC492788	insulin receptor; insulin receptor (AA at 78); insulin receptor precursor
THC492789	insulin receptor precursor
THC496276	IGF-BP 4; insulin-like growth factor binding protein 6; insulin-like growth factor binding protein 4
THC497182	IGF-I receptor; insulin-like growth factor I receptor; insulin-like growth factor 1 receptor precursor
THC498325	insulin induced protein 2
THC503886	JAK3M; INL3_HUMAN [AA 1- 65]; Leydig insulin-like peptide; Ley I-L=Leydig insulin-like peptide [human, testis, Peptide, 131 aa]; INL3_HUMAN; LEY-I-L; RELAXIN-LIKE FACTOR; RLF
THC503887	insulin-like leydig hormone; Leydig insulin-like hormone; insulin-like 3 (Leydig cell)
THC506136	sulfonylurea receptor; sulfonylurea receptor 1; alternative splice (exon 17); sulfonylurea receptor (hyperinsulinemia); Homo sapiens sulfonylurea receptor (SUR1) gene, exon 39
THC506396	insulin-degrading enzyme
THC509257	ribosomal protein S15; human homologue of rat insulinoma gene (rig); putative; rig-analog protein (putative); putative
THC511523	insulin receptor substrate 4
THC512701	Cyr61 protein; Gig1 protein; tumor RMS cell line RD specific product; hCYR61 protein; growth-factor inducible immediate early gene product CYR61; insulin-like growth factor binding protein 10
THC512702	CYR61 PROTEIN PRECURSOR (GIG1 PROTEIN) (INSULIN-LIKE GROWTH FACTOR- BINDING PROTEIN 10). GP|2832337|
THC512703	CYR61 PROTEIN PRECURSOR (GIG1 PROTEIN) (INSULIN-LIKE GROWTH FACTOR- BINDING PROTEIN 10). GP|2832337|
THC517432	bG256O22.1 (similar to IGFALS (insulin-like growth factor binding protein, acid labile subunit))
THC517523	mannose 6-phosphate/insulin-like growth factor-II receptor
THC521405	zinc finger-like protein IA-1; insulinoma-associated 1
THC524839	insulin receptor substrate 1; hIRS-1=rat insulin receptor substrate-1 homolog [human, cell line FOCUS, Peptide, 1243 aa]; insulin receptor substrate-1, IRS-1 [human, skeletal muscle, Peptide, 1242 aa]; insulin receptor substrate 1; IRS-1
THC526218	insulin receptor-related receptor
THC526500	mannose-6-phosphate receptor, cation independent; mannose 6-phosphate receptor; insulin-like growth factor type II receptor; insulin-like growth factor II receptor; insulin-like growth factor 2 receptor; dJ249F5.1 (insulin-like growth factor 2 receptor);
THC529746	preprotein 12 (AA -25 to 234); PP12 precorsor (AA -25 to 234); small IGF-binding-protein; insulin-like growth factor binding protein precursor; insulin-like growth factor binding protein 1; insulin-like growth factor binding protein-1; IGF-binding preprot
THC535179	BETA3=class B basic helix-loop-helix transcription factor {hamsters, HIT insulin tumor cells, Peptid
THC540587	insulin receptor tyrosine kinase substrate; similar to insulin receptor substrate BAP2; similar to PID:g4126477
THC541846	insulin-like 6; relaxin/insulin-like factor 1; insulin-like protein 6
THC543189	insulin promoter factor 1; insulin promoter factor 1 (IPF-1); insulin promoter factor 1, homeodomain transcription factor; insulin promoter factor 1, IPF-1 [human, Peptide, 283 aa]
THC544492	insulin-like growth factor binding protein 2; insulin-like growth factor binding protein 2, 36 kDa; precursor polypeptide (AA -39 to 289); insulin-like growth factor binding protein-2, IGFBP-2 [human, placenta, Peptide, 328 aa]
THC546425	CATION-INDEPENDENT MANNOSE-6-PHOSPHATE RECEPTOR PRECURSOR (CI MAN-6-P RECEPTOR) (CI-MPR) (INSULIN-LI

(Table 1) contd.....

Sequence	Putative ID
THC550306	insulin-like growth factor binding protein complex; acid-labile subunit; insulin-like growth factor binding protein, acid-labile subunit; insulin-like growth factor binding protein, acid labile subunit
THC554077	insulin induced gene 1; insulin induced protein 1
THC558697	insulin receptor substrate-2; HSPC298; insulin receptor substrate 2
THC558800	prolyl 4-hydroxylase beta subunit (AA 1-491); prolyl 4-hydroxylase beta-subunit; prolyl 4-hydroxylase, beta subunit/protein disulfide isomerase/thyroid hormone-binding protein; glutathione-insulin transhydrogenase (216 AA); procollagen-proline, 2-oxogluta
THC560652	insulin; pre-proinsulin; preproinsulin; proinsulin precursor
THC561837	insulin-like peptide INSL5; insulin-like 5
THC562276	insulin upstream factor 1
THC563088	insulin activator factor; insulin activator factor, INSAF [human, Pancreatic insulinoma, Peptide Partial, 744 aa]
HT3739	insulin-like growth factor 2
NP078669	[GB:S51971] insulin-like growth factor-II
NP084409	[GB:AF064078] insulin receptor-related receptor
NP103481	[GB:S73149] orf in intron 7 of insulin-like growth factor II gene
NP104342	[GB:AF053356] insulin receptor substrate like protein
NP176148	[GB:AL022322] bK228A9.2 (novel protein similar to FAS-ligand associated factor 3 and Insulin receptor tyrosine kinase 53 kD substrate)

The human insulin genomic sequence map (with exons, introns and approximate restrictions cut sites) and its cDNA (equivalent to the cytoplasmic mRNA for insulin its and its restriction product).

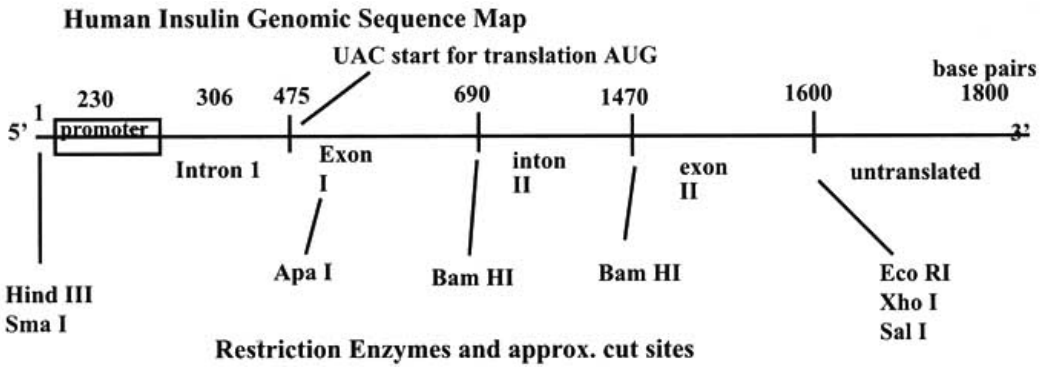


Fig. (2). Human Insulin genomic map.

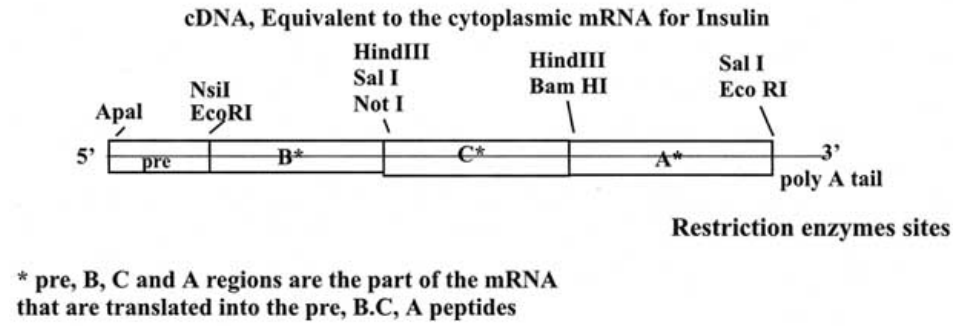


Fig. (3). cDNA i.e. equivalent to the cytoplasmic mRNA insulin.

THE AMINO ACID SEQUENCE OF PORCINE INSULIN

The complete amino acid sequence of porcine insulin was determined by Sanger and co- workers in 1955 [1]. Insulin has a molecular weight of 5808 daltons and contains 51 amino acids, arranged in two chains (an acidic A chain of 21 residues and a basic B chain of 30 residues) which are cross-linked by two disulphide bridges. A third disulphide bridge links two parts of the shorter A chain [16].

THE CRYSTAL STRUCTURE OF PORCINE INSULIN

Although insulin was first crystallized by Abel and co-workers in 1925, it was not until 1934 that Scott discovered that the rhombohedral crystals were a zinc-insulin complex. Together with Fischer [17], Scott showed that the zinc might be replaced by other divalent ions. In time, this opened up the possiblity that the crystal structure might be solved by isomorphous replacement. An initial x-ray analysis of porcine zinc insulin at 2.8 Å was completed in 1969 by Hodgkin and co-workers and at 2.5Å by the Peking Insulin Structure Research Group in 1971[18,19]. 2-zinc insulin has since been analysed at 1.9 Å resolution, at 1.8Å resolution by Peking Insulin Structure Research Group in 1974 [20] and more recently refined to 1.5 Å [21,22] and to 1.1 Å resolution [23]. In the rhombohedral crystal, the unit cell contains six molecules of insulin. The hexamer is assembled from three equivalent dimers which are related by a 3-fold axis. Two zinc ions lie 17 Å apart on the 3-fold axis, each of them co-ordinated by B10 His from each of the 3 dimers and 3 water molecules.

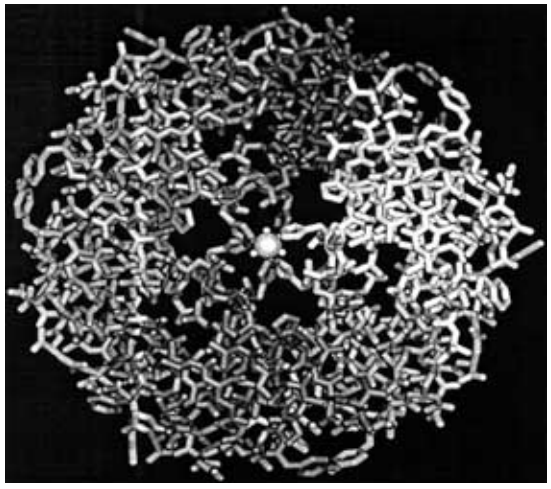


Fig. (4). A view of the hexamer perpendicular to the 3-fold axis.

The insulin hexamer viewed down the 3-fold axis, indicates the position of the zinc ions and the three water molecules that coordinate with each of them.

THE INSULIN MONOMER

The insulin monomer is a compact globular structure with a hydrophobic core. Although the surface residues are primarily polar, there are two hydrophobic surfaces on each side of the molecule that are buried during the formation of dimers and 2-zinc hexamers. In the insulin fold, the A chain is a compact unit around which the B chain is wrapped.

The A chain consists of two anti-parallel stretches of imperfect alpha helices (A2 Ile - A8 Thr and A13 Leu - A19 Tyr) which are joined by a turn from A9 Ser to A12 Ser, stabilized by the A6-A11 disulphide. The A chain lies in a plane in which the N and C terminii are brought to the same side, bringing A2 Ile and A19 Tyr into van der Waals contact.

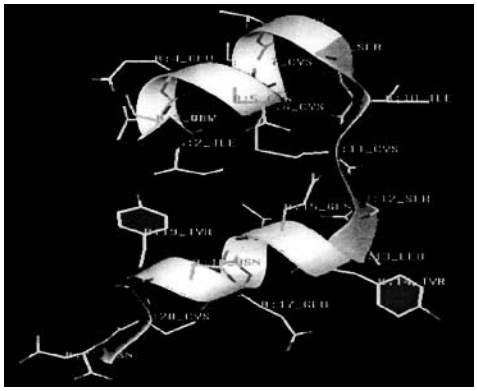


Fig. (5). The A chain lies in a plane in which the N and C terminii are brought to the same side, bringing A2 Ile and A19 Tyr into van der Waals contact.

The B chain consists of an alpha-helix (B9 Ser - B19 Cys) from which both N and C terminii residues extend. The glycine residues at B20 and B23 allow the chain to fold back on itself in an approximate V- shape, and this brings the C terminal residues B24 Phe and B26 Tyr into van der waals contact with B15 Leu and B11 Leu of the alpha-helix.

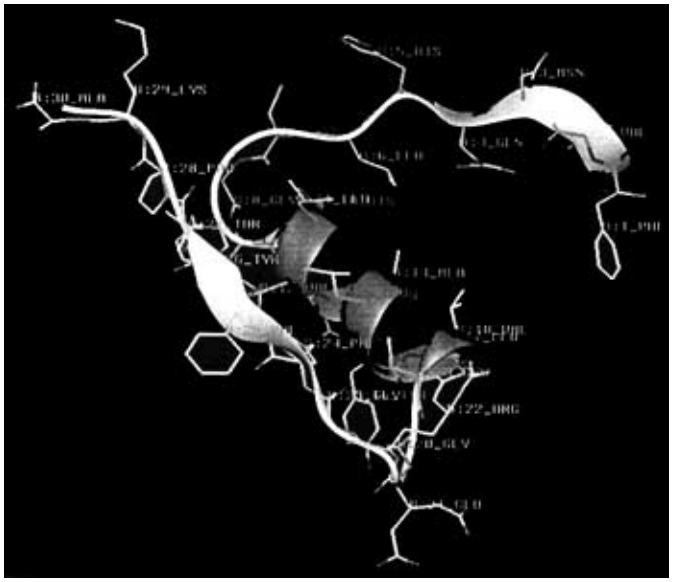


Fig. (6). The glycine residues at B20 and B23 allow the chain to fold back on itself in an approximate V- shape, and this brings the C terminal residues B24 Phe and B26 Tyr into van der waals contact with B15 Leu and B11 Leu of the alpha-helix.

The insulin fold is formed when interchain disulphides at A7 and A20 form interchain disulphides with the B chain cysteines at B7 and B19 respectively. The (A7-B7) disulphide is fully exposed on the surface of the molecule, whereas the (A20-B19) disulphide is part of the hydrophobic core. Burial of the intrachain (A6 - A11) disulphide and the non-polar side chains of A16 Leu, B11 Leu, B15 Leu, A2 Ile and B24 Phe provides the hydrophobic interior stabilising the fold. The N termini residues of the B chain are folded across and run anti-parallel to the turn in the A chain, giving rise to hydrogen bonding between A11 Leu and B4 Gln, A7 Cys and B5 His, and between A19 Tyr and B25 Phe. Further stability is also provided by a salt bridge between the polypeptide chains at A11 Cys and B4 Gln, between the A7 carbonyl oxygen and the B5 His side chain, and between the A19 carbonyl oxygen and the B25 back-bone nitrogen. Further stability is also provided by a salt bridge between B29 Lys and A4 Glu and between the positively charged B22 Arg side chain and the negatively charged A21 terminal carboxyl group.

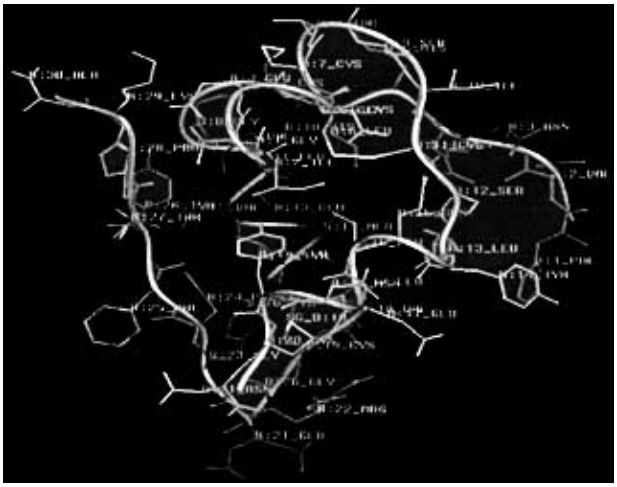


Fig. (7). The insulin monomer viewed perpendicular to the 3-fold axis. The 2 interchain and 1 intrachain disulphides are shown in yellow.

THE AGGREGATION SURFACES OF INSULIN

In solution at neutral pH and at physiological concentrations (about 1 ng/ml) insulin exists as a monomer [24], and it is the monomer which is the active form of the hormone [25]. At higher concentrations at acid or neutral pH (in the absence of zinc) the insulin monomer self-associates to form dimers and (in the presence of zinc) hexamers [26, 27]. Studies of insulin analogues demonstrate how activity depends on the integrity of the insulin fold, and also allow mapping of the interactive residues on the surface of the molecule. The central residues responsible for hexamer formation include B10 His, which binds zinc ons, and B14 Ala, B17 Leu, B20 Gly and A 13 Leu, which are involved in close-packed hydrophobic interactions. The ability to form dimers is mediated by hydrophobic interactions involving B8 Gly, B9 Ser, B12 Val, B13 Glu, B16 Tyr, B24 Phe, B25 Phe, B26 Tyr, B27 Thr and B28 Pro [28,29]. The B chain of one monomer packs against the B chain of

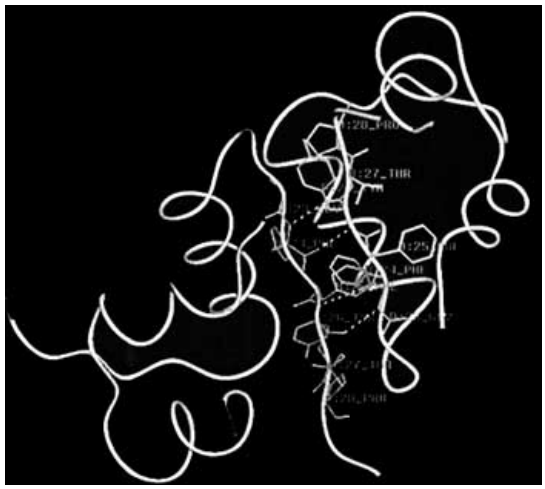


Fig. (8). The porcine insulin dimer.

the second monomer and further stability is provided by hydrogen bonding between the antiparallel beta-strands B24 - B26 of each molecule. The packing arrangement in the dimer results in perturbations of side-chain and main-chain structure [30], such that the monomers differ slightly in conformation.

CONFORMATIONAL VARIATION IN THE INSULIN MONOMER

The conformation of the insulin monomer varies from one crystal form to another [31]. Although the insulin-fold is always retained, the major secondary structure elements are capable of individual rigid body movements that are accomodated by alterations in side-chain positions [32]. The variety of insulin quaternary structures have been the subject of much discussion, but the differences are mainly localised to the N- and C- terminii of the B-chain [33]. The most extreme variation is in the N-terminus of the B chain which may exist either in an extended conformation (**T** state) or as an extension of the B chain helix (**R** state) [34,35].

That these conformational variations are not simply due to crystal packing forces is demonstrated by molecular dynamics simulations performed *in vitro* which highlight relative movement of the helices and flexibility of the chain termini [36,37]. The **T** to **R** transition can be reproduced reversibly in the restricting environment of the crystal lattice, demonstrating the flexibility inherent in the insulin molecule [38]. The **T** to **R** transition does not appear to be important in receptor binding since gamma-ethylnorleucine B6 insulin, which has limited ability to undergo the transition, still binds the insulin receptor with high affinity [39].

INSULIN X-RAY STRUCTURES

Since the initial X-ray analysis of insulin [18], there have been a series of crystallographic studies on monomeric, dimeric and hexameric insulins in several forms. Some of

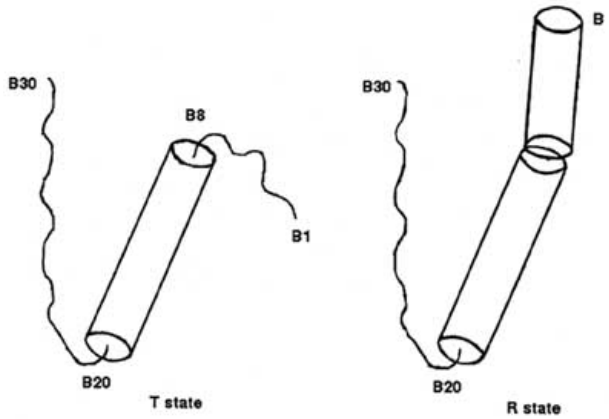
these insulins have been variants and others have been chemically synthesized, modified or mutated by recombinant DNA techniques.

Although there are variations in the molecule's conformation in the different crystals and between species, the insulin fold is essentially maintained. As the native insulin monomer will always form multimers at the concentrations necessary for crystal growth, a crystal structure analysis of the native monomeric molecule is impossible. However, despentapeptide insulin which is unable to form the hydrogen bonded beta-sheet between the monomers , does crystallize as a monomer (although there are still 2 molecules in the unit cell) and gives monoclinic crystals [40,41]. In the absence of metal ions native insulin crystallizes as a dimer. In some crystals the two-fold axis is exact and crystallographic e.g. neutral pH cubic crystals [30] and neutral pH tetragonal crystals [42] and the insulin monomer occupies the asymmetric unit. In other crystals the two-fold axis is approximate and non-crystallographic e.g. low pH orthorhombic human insulin crystals [43]. The two B25 Phe side-chains respond to the symmetry of the dimer; in a symmetrical dimer both are directed away from the two-fold axis, while in the asymmetrical dimer one side turns away and contacts its partner across the approximate axis. The conformation of the insulin molecule which persists in nearly all crystal forms and probably also in the circulation. In the dimerization of two insulin monomers the two B13 Glu side chains are directed away from each other and solvated, minimizing their repulsion [44]. In hagfish insulin, which has an Asn residue at B13, the two molecules of the dimer are identical and resemble molecule 2 of the porcine dimer [45,46].

In the presence of zinc ions, three identical insulin dimers assemble into a hexamer in which each of two zinc ions lies on a three-fold axis and resides in a distorted octahedral ligand field co-ordinated by three B10 His imidazolyl groups and by three water molecules. The hexamer is stabilized by hydrophobic interactions between B1 Phe, B14 Ala, B17 Leu, A13 Leu and B18 Val residues. The local two-fold axis within each dimer is perpendicular (or approximately so) and intersects the three-fold axis. X-ray crystallographic studies of porcine 2 zinc insulin [20] showed that the hexamer is torus shaped (50 Å in diameter by 35 Å high) and that the two zinc ions lie 17 Å apart on either side of the hexamer centre, on the three-fold axis which traverses the central cavity. In the porcine 2 Zn insulin crystal (T6) the dimer is the asymmetric unit [28].

The insulin hexamer is toroidal, highly stable and has a polar surface, making it ideal for crystal formation. In pancreatic beta-cells (which have a high content of zinc), mature secretory granules contain insulin which forms a dense array of 50 Å units with packing similar to that found in the crystalline hexamer [47], and it has been proposed that the 2 Zn hexamer is the storage form of insulin [20]. Crystallization within the storage granule seems to serve two purposes, firstly to act as an efficient store of insulin and secondly to protect insulin molecules from proteolytic processing.

It has been shown that anions induce the **T** to **R** transition at residues B1-B8 of the insulin molecule [48,49,50]. Transformation efficiency follows the Hoffmeister lyotropic series [51,52] and some residues move by as much as 20-25 Å [53,54]. Chloride ion concentrations above 6% have a dramatic effect on the insulin hexamer, inducing the **T** to **R** transition in three molecules , which now form a three-fold symmetrical trimer, out of the six, and causing a cleft to appear in these molecules due to movement of the (A7-B7) disulphide. The movement apart of the A-chain and the B-chain C-terminii breaks the hydrogen bond between A19 Tyr and B25 Phe and exposes the hydrophobic core residues. The approximate two- fold symmetry between the dimers in the 2 Zn hexamer is lost, however the axial symmetry relating the monomers in the dimer is preserved. At one face of the hexamer, a single octahedral zinc site coordinated by three B10 His residues and three water molecules is retained. At the other face of the hexamer the helical conformation of the B1-B8 segments of the other three B-chains brings the B5 His and B10 His rings from adjacent dimers together, creating two types of additional zinc sites. At these new faces, the B10 His side chain can exist in two distinct orientations. In one orientation the symmetrically related B10 His side chains



The T→R transition in insulin (after Kaarsholm *et al*, 1989)

Fig. (9). The T → R transition in insulin.

and a chloride ion (or water molecule) coordinate a zinc ion on the 3-fold axis. Alternatively, the B10 His side chains may be rotated about the C-alpha-C-beta bond to bring them into close proximity with the imidazole ring of B5 His from the adjacent dimer, creating three identical off-axial sites. Each off-axial zinc ion is coordinated by two chloride ions (or water molecules) and two histidyls (B5 His and B10 His). The off-axial sites are not always fully occupied [53] but at high zinc concentrations 4 Zn hexamers can be formed [55]. In the 4 Zn (T3,R3) porcine and human insulin crystals, the dimer is still the asymmetric unit [56,54]. The 2 Zn to 4 Zn crystal transformation can be elicited by soaking the crystalline 2 Zn hexamer in high concentrations of anions without distorting the crystal lattice. If the chloride ion concentration is reduced below 6%, the 4 Zn (T3,R3) crystal reverts back to the 2 Zn (T6) structure [38]. Alternatively, the 4 Zn hexamer can be crystallised directly.

Phenol has also been shown to induce the 2 Zn to 4 Zn transformation in solution [57], but 2Zn insulin crystals do not survive soaking in phenol without cracking. However, crystallization in the presence of phenol at levels of 0.5%, causes the B-chain helices of all insulin molecules to extend from B1 to B20, giving the dimers approximate two-fold symmetry. Each phenol binds in a hydrophobic cavity created by the packing of the B1 to B8 helix against the A chain of the three-fold related dimer. More specifically, the phenol hydroxyl group hydrogen bonds to the (A6-A11) cysteine bridge and the aromatic ring makes van der Waals contact with B5 His of the adjacent dimer, blocking off the axial zinc-binding sites. The resulting 2 Zn (R6) crystals are monoclinic with a hexamer in the asymmetric unit [58]. Phenol is commonly used as a preservative in commercial insulin preparations at a level of approximately 0.2%, and interestingly such levels can occasionally convert rhombohedral crystals to the monoclinic form [59].

When insulin is mixed with protamine in the presence of zinc, phenol or *m*cresol, a precipitate is formed which can be converted to tetragonal crystals [60]. Chemical analysis indicates the presence of two zinc ions and 22 phenol molecules per hexamer and one protamine per 8.5 insulin monomers. X-ray studies reveal that the hexamer in these crystals is symmetric, with three two-fold axes as is the case for monoclinic crystals [54].

The reluctance of native insulin to form hexamers in the absence of zinc ions suggests that there are unfavourable interactions which are overcome by the zinc coordination to B10 His. The repulsive potential of the six B13 Glu carboxylate groups at the centre of the insulin hexamer has been predicted to limit the formation of zinc-free insulin hexamers at neutral pH [20,28,61]. In the 2 Zn hexamer the B13 Glu side chains are arranged as three hydrogen-bonded pairs, an arrangement which reduces their electrostatic repulsion at the expense of the energy required to protonate half of the carboxylates. Solution studies including osmometry [62,63], volume exclusion chromatography and circular dichroism studies [64] have established the existence of a stable metal-free hexamer when B13 Glu is mutated to Gln. X-ray analysis of the zinc-free B13 Gln insulin hexamer reveals that the glutamine side-chains are entirely hydrated and make no hydrogen-bonding contacts [65]. Surprisingly the organisation of the zinc-free hexamer is T3,R3 in the absence of any anions; either the B13 Gln mutation favours the **R** state or the sulphate ion stabilizes the B1 - B8 helix (ammonium sulphate is not used for 2 Zn hexamer growth as a change in pH is sufficient to cause supersaturation). Inspection of known insulin sequences shows that a carboxylic acid residue (Glu or Asp) at B13 is almost universal in all zinc binding insulins [66], the exception being the unusual Atlantic Hagfish insulin where B13 is an asparagine and which is thought not to form hexamers [38,42], suggesting that it has important structural or functional properties. It has been proposed that the function of B13 acidic groups is to drive apart the hexamer as soon as the stabiliizing zinc ions diffuse away upon expulsion of the hexamer from the beta-cell [65]. However, B13 Glu is conserved in Guinea Pig insulin which is known not to hexamerise [67], suggesting that its role may not be limited to hexamer assembly. Commensurate with this is the 80% fall in potency of B13 Gln insulin in the free fat cell assay and a fall of 60% in the mouse blood glucose assay [68].

The work of Schlichtkrull [55] on the stoichiometries of divalent metal ion binding in crystalline rhombohedral insulin hexamers gave the first indication for the presence of a third, high affinity metal binding site within the insulin hexamer. During the initial X-ray structure determination of the 2 Zn hexamer, isomorphous replacement studies located the extra metal binding site to the centre of the hexamer [69] and further studies showed that the B13 Glu residues specifically bound metal [20,70]. NMR experiments involving metal ion substitution established that the 2 Zn hexamer is a calcium binding protein [71]. A variety of physico-biochemical studies have since shown that one calcium ion binds at the centre of each hexamer in a cage formed by the six B13 Glu carboxylates with a dissociation constant of 80 microM [71-74]. Substitution of a calcium ion into the metal-free B13 Glu cavity disrupts the intra-dimer Glu-Glu hydrogen-bonding interaction, and the single ion has to move rapidly between the three symmetry related sites in order to coordinate the covalently seperate molecules. In this respect the calcium binding site of the 2 Zn hexamer is quite unlike any other known protein structure [75]. The presence of calcium at the centre of the hexamer has a stabilising effect by eliminating the coulombic effects whih would otherwise arise from ionisation of the B13 carboxyls in its absence. It has been shown that pancreatic storage granules contain high concentrations of Ca2+ in addition to Zn2+ [76], and their physiological role is probably to facilitate hexamer formation from proinsulin and stabilization of the complex during storage.

Several mutant insulins have been crystallised with recipies that differ from those traditionally used, probably owing to their altered stability and solubility [77].

RECOMBINANT DNA TECHNOLOGY IN THE SYNTHESIS OF HUMAN INSULIN IN *ESCHERICHIA COLI*

A strain of the common bacterium, *Escherrichia coli (E. coli)* , an inhabitant of the human digestive tract, is the

'factory' used in the genetic engineering of insulin. When the bacterium reproduces, the insulin gene is replicated along with the plasmid, a circular section of DNA . *E. coli* produces enzymes that rapidly degrade foreign proteins such as insulin. By using mutant strains that lack these enzymes, the problem is avoided.

To make the bacteria produce insulin, the insulin gene needs to be tied to this enzyme. Restriction enzymes, naturally produced by bacteria, act like biological scalpels, only recognizing particular stretches of nucleotides, such as the one that codes for insulin. This makes it possible to sever certain nitrogen base pairs and remove the section of insulin coding DNA from one organism's chromosome so that it can manufacture insulin. DNA ligase is an enzyme that serves as genetic glue, welding the sticky ends of exposed nucleotides together.

The first step is to chemically synthesise the DNA chains that carry the specific nucleotide sequences characterizing the A and B polypeptide chains of insulin .

The required DNA sequence can be determined because the amino acid compositions of both chains have been charted. Sixty three nucleotides are required for synthesizing the A chain and ninety for the B chain, plus a codon at the end of each chain, signaling the termination of protein synthesis. An anti-codon, incorporating the amino acid, methionine, is then placed at the beginning of each chain which allows the removal of the insulin protein from the bacterial cell's amino acids. The synthetic A and B chain 'genes' are then separately inserted into the vector(may be plasmid).

The recombinant plasmids are then introduced into *E. coli* cells. Practical use of Recombinant DNA technology in the synthesis of human insulin requires millions of copies of the bacteria whose plasmid has been combined with the insulin gene in order to yield insulin. The protein which is formed, joined to either the A or B chain of insulin . The A and B chains are then extracted purified.

The two chains are mixed and reconnected in a reaction that forms the disulfide cross bridges, resulting in pure synthetic human insulin .

Human insulin is the only animal protein to have been made in bacteria in such a way that its structure is absolutely identical to that of the natural molecule. This reduces the possibility of complications resulting from antibody production. In chemical and pharmacological studies, commercially available Recombinant DNA human insulin has proven indistinguishable from pancreatic human insulin. Initially the major difficulty encountered was the contamination of the final product by the host cells, increasing the risk of contamination in the fermentation broth. This danger was eradicated by the introduction of purification processes. When the final insulin product is subjected to a battery of tests, including the finest radio-immuno assay techniques, no impurities can be detected. The entire procedure is now performed using yeast cells as a growth medium, as they secrete an almost complete human insulin molecule with perfect three dimensional structure. This minimises the need for complex and costly purification procedures.

Goeddel *et al.*, [5] and Wong [6] illustrate the strategies in their papers. The A and B chains of the insulin gene are synthesized chemically and that is controlled by the *lac* promoter, which is inducible by IPTG and/or X-gal (the substrate). The construction is such that the insulin chains would be made fusion proteins joined by a methionine to the end protein. Then, the insulin chains (either A-chain or B-chain) can be cleaved off from at the methionine residue by cyanogen bromide (CNBr) treatment. The transformants will be selected by using ampicillin resistant gene marker on the expression vector. In the original work, pBR322 is used for as cloning vector [3]. At present, the new generation plasmid such as pUC18 and pUC19 may be work equally well or better, as pUC18 or pUC19 has higher copy numbers per cell than its ancestor pBR322 does. The A-chain and B-chain products can be purified separately before disulfide bonds through the reduction-reoxidation reactions join them. The correct joining and folding will become active mature insulin ready to use.

FIRST EVIDENCES OF SUCCESSFUL EXPRESSION OF HUMAN INSULIN IN *ESCHERICHIA COLI*

As commercialization process is difficult only three companies in the world have recombinant insulin. In the original work, Goeddel and his colleagues [5] used the radioimmunoassay kit from Pharmacia (Phadebus Insulin Test). The sequences of DNA were obtained after cloning and plasmid construction could be directly verified as designed. The *E. coli* products, after cyanogen bromide cleavage, behaved as insulin chains in three different principles, i.e., gel filtration, ion exchange, and reverse-phase high-performance liquid chromatography (HPLC). The scanning and transmission electron micrographs showed the inclusion bodies accumulated in the bacterial cells due to the production of the gene product. Now several scientists have reported successful expression of human insulin in *Escherichia coli* from different laboratories.

CONCLUSION

When using *Escherichia coli* or bacteria as a host for transformation for human protein production, problems might be encountered but there are solutions. [78]. For example, bacterial RNA polymerases do not recognize eukaryotic promoters. Thus, the human gene must be placed adjacent to a strong bacterial promoter which is associated with a ribosome binding site and an initiating ATG codon. Another problem is that bacteria do not have post-trancriptional modification such as mRNA splicing (cleaving off introns and joining exons). This problem can be solved by using cDNA (synthesized from mature mRNA by reverse transcriptase enzyme) or by chemically synthesizing the gene if the gene is short, e.g., somatostatin gene which is only 51 bp long and encodes for 14 amino acids in the sequence. Sometimes bacteria do not have post-translational

modification such as glycosylation (addition and attachment of sugar groups). This problem can be eliminated by using yeast as host, since yeast is eukaryotic allowing this process to occur. This low cost recombinant DNA insulin can fulfill the total insulin demand of the world in future.

REFERENCES

[1] Brown, T.A. *Gene Cloning*, 3rd edition. Chapman & Hall. London, UK,**1995**.

[2] Itakura, K.; Hirose, T.; Crea, R. *Science*, **1977**, *198*, 1056-1063.

[3] Goeddel, D.V.; Kleid, D.G.; Bolivar, F. *Proc. Natl. Acad. Sci. USA*, **1979,** *76*, 106-110.

[4] Hitzeman, R.A.; Hagie, F.E.; Levine, H.L. *Nature*, **1981**, *293*, 717-722.

[5] Goeddel, D.V.; Heyneker, H.L.; Hozumi, T. *Nature,* **1979**, *81*, 544-549.

[6] Wong, W.S. *The ABCs of Gene Cloning.* Chapman & Hall. New York, NY, USA, **1997** .

[7] Georges, F.; Brousseau, R.; Michniewicz, J. *Gene,* **1984**, *27*, 201-211.

[8] Bell, G.I.; Selby, M.J.; Rutter, W.J. *Nature,* **1982**, *5844*, 31-35.

[9] Ullrich, A.; Dull, T.J.; Gray, A. *Science,* **1980**, *209*, 612-615.

[10] Ullrich, A.; Dull, T.J.; Gray, A. *Nucleic Acids Research,* **1982**, *10*, 2225-2240.

[11] Bell, G.I.; Swain, W.F.; Pictet, R. *Nature,* **1979**,5738, 525-527.

[12] Bell, G.I.; Pictet, R.L.; Rutter, W.J. *Nature,* **1980**,*5751*, 26-32.

[13] Sures, I.; Goeddel, D.V.; Gray, A.; Ullrich, A. *Science*, **1980**, *208*,57-59.

[14] Brousseau, R.; Scarpulla, R.C.; Sung, W. *Gene,* **1982**,*17*, 279-289.

[15] Narnag, S.A.; Brousseau, R.; Georges, F. *Can. J. Biochem.,* **1984**, *62*, 209-216.

[16] Ryle, A.P.; Sanger, F.; Smith, L.F.; Kitai, R. *Biochemical J.,* **1955**, *60*, 541.

[17] Scott, D.A.; Fisher, A.M. *J. Pharmacol. Exp. Ther.,* **1934**, *55*, 206.

[18] Adams, M.J.; Blundell, T.L.; Dodson, E.J.; Dodson, G.G.; Vijayan, M.;Baker,E.N.;Harding, M.M.; Hodgkin, D.C.; Rimmer, B.; Sheat, S. *Nature*,**1969**, *224,* 491.

[19] Blundell, T.L.; Cutfield, J.F.; Cutfield, S.M.; Dodson, E.J.; Dodson, G.G.; Hodgkin, D.C.; Mercola, D.A.; Vijayan, M. *Nature,* **1971**, *231*, 506.

[20] Blundell, T.L.; Dodson, G.G.; Hodgkin, D.C.; Mercola, D.A. *Adv. Prot. Chem.,* **1972**, *26*, 279.

[21] Sakabe, N.; Sakabe, K.; Saski, K. *Diabetes,* **1977**, *5*, 12.

[22] Dodson, G.G.; Cutfield, S.; Hoenjet, E.; Wollmer, A.; Brandenburg, D.D. In *Insulin, Chemistry, Structure and Function of Insulin and Related Hormones*. Brandenburg, D.; Wollmer, A. Eds.; Walter de Gruyter, New York, **1980**,pp.17-20.

[23] Sakabe, N.; Sakabe, K.; Saski, K. In *Structural Studies of Molecules of Biological Interest.* Dodson,G.G.; Glusker, J.P. Eds.; Oxford, Clarendon Press, **1981**, pp.509.

[24] Frank, B.H.; Pekar, A.H.; Veros, A.J. *Diabetes Supp.,* **1972,** *21*, 486.

[25] Cahill, G.F. *Diabetes,* **1971**, *20*, 780-797.

[26] Jeffrey, P.D.; Coates, J.H. *Biochemistry,* **1966**, *5*, 4889.

[27] Carpenter, F.H. *Am. J. Med.,* **1966**, *40*, 750.

[28] Baker, E.N.; Blundell, T.L.; Cutfield, J.F.; Cutfield, S.M.; Dodson, E.J.; Dodson, G.G.; Hodgkin, D.C.; Hubbard, R.E.; Issaacs, N.W.; Reynolds, C.D.; Sakabe, K. *Philos. Trans. R. Soc. London,* **1988**, *B319*, 369.

[29] Brange, J.; Owens, D.; Kang, S.; Volund, A. *Diabetes Care,* **1990**, *13*, 923.

[30] Dodson, E.J.; Dodson, G.G.; Lewitova, A.; Sabesan, M. *J. Mol. Biol.,* **1978**, *125*, 387.

[31] Derewenda, U.; Derewenda, Z.S.; Dodson, G.G.; Hubbard, R.E. In *Insulin Structure ; Handbook of Experimental Pharmacology,* Cautrecasas, P.; Jacobs, S. Eds.; Springer-Verlag , **1990**,*92*, pp. 23.

[32] Lesk, A.M. In *Protein Architecture, A Practical Approach* , Oxford University Press, **1991**, pp.135.

[33] Dodson, G.G. Cutfield, S., Hoenjet, E., Wollmer, A. & Brandenburg, D.D. In *Insulin, Chemistry, Structure and Function of Insulin and Related Hormones* Brandenburg, D.; Wollmer, A. Eds.; Walter de Gruyter, New York, **1980**,17.

[34] Kaarlsholm, N.C.; Ko, H.-C.; Dunn, M.F. *Biochemistry,* **1989**, *28*, 4427.

[35] Kruger, P.; Gilge, G.; Cabuk, Y.; Wollmer, A. *Biol. Chem.,* **1990***, 371*, 669.

[36] Kruger, P.; Stassburger, W.; Wollmer, A.; van Gunsteren, W.F.; Dodson,G.G. *Eur. Biophys. J.,* **1987**, *14*, 449.

[37] Caves, L.S.D.; Nguyen, D.T.; Hubbard, R.E. In *Molecular Dynamics : Applications in Molecular Biology* . Ed. J.M.Goodfellow, Macmillan, London, **1990**.

[38] Bentley, G.A.; Dodson, G.G.; Lewitova, A. *J. Mol. Biol.,* **1978**, *126*, 871.

[39] Nakagawa, S.H.; Tager, H.S. *J. Biol. Chem.,* **1991**, *266*, 11502.

[40] Bi, R.C.; Dauter, Z.; Dodson, E.J.; Dodson, G.G.; Gordino, F.; Hubbard, R.; Reynolds, C.D. *Proc. Indian Acad. Sci. (Chem.Sci.),* **1983**, *92*, 473.

[41] Bi, R.C.; Dauter, Z.; Dodson, E.J.; Dodson, G.G.; Giordano, F.; Reynolds, C. *Biopolymers,* **1984**, *23*, 391.

[42] Cutfield, J.F.; Cutfield, S.M.; Dodson, E.J.; Dodson,G.G.; Emdin, S.F.; Reynolds, C.D. *J. Mol. Biol.,* **1979**, *132*, 85.

[43] Derewenda, U.; Derewenda, Z.; Dodson, E.J.; Dodson, G.G.; Reynolds, C.D.; Smith,G.D.; Sparks, C.; Swenson, D. *Nature,* **1989**, *388*, 594.

[44] Badger, J.; Dodson, E.J.; Dodson, G.G.; Harris, M. *Acta. Crystallogr.,***1991**,*47*, 127.

[45] Cutfield, J.F.; Cutfield, S.M.; Dodson, E.J.; Dodson, G.G.; Sabesan, M.N. *J. Mol. Biol.,* **1974**, *87*, 23.

[46] Peterson, J.D.; Steiner, D.F.; Emdin, S.O.; Falkmer, S. *J. Biol. Chem.,* **1975**, *250*, 5183.

[47] Greider, M.H.; Howell, S.L.; Lacy, P.E. *J. Cell Biol.,* **1969**, *41*, 162.

[48.] Bentley, G.A.; Dodson, G.G.;Dodson, E.J.; Hodgkin, D.C.; Mercola, D.A.; Wollmer, A. *Spring Meeting of the British Diabetic Association*, Sheffield, U.K, **1975**, 122.

[49] Renscheidt, H.; Strassburger, W.; Glatter, U.; Wollmer, A.; Dodson, G.G.; Mercola, D.A. *Eur. J. Biochem.,* **1984**, *142*, 7.

[50] Kaarlsholm, N.C.; Dunn, M.F. *Biochemistry,* **1987**, *26*, 883.

[51] Harding, M.M.; Hodgkin, D.C.; Kennedy, A.F.; O'Connor, A.; Weltzmann, P.D.J. *J. Mol. Biol*, **1966**, *16*, 212.

[52] De Graaf, R.A.G.; Lewit-Bentley, A.; Tolley, S.P. In *Structural Studies on Molecules of Biological Interest,* Dodson, G.G.; Glusker, J.P.; Sayre, D. Eds., Clarendon Press, Oxford, **1981**, 547.

[53] Smith, G.D.; Swenson, D.C.; Dodson, E.J.; Dodson, G.G.; Reynolds, D. *Proc. Natl. Acad. Sci. USA,* **1984**, *81*, 7093.

[54] Derewenda, U.; Derewenda, Z.S.; Dodson, G.G.; Hubbard, R.E.; Korber, F. *Br. Med. Bull.,* **1989**, *45*, 19.

[55] Schlichtkrull, J. *Insulin Crystals*, Copenhagen, Munksgaard, **1958**, *2*, 205.

[56] Bentley, G.A.; Dodson, E.J.; Dodson, G.G.; Hodgkin, D.C.; Mercola, D.A. *Nature,* **1976**, *261*, 166.
[57] Wollmer, A.; Rannefield, B.; Johansen, B.R.; Hejnaes, K.R.; Balschmidt, P.; Hansen, F.B. *Hoppe-Seyler's Z Biol. Chem.,* **1987**, *368*, 903.
[58] Derewenda, U.; Derewenda, Z.; Dodson, E.J.; Dodson, G.G.; Reynolds, C.D.; Smith,G.D.; Sparks, C.; Swenson, D. *Nature,* **1989**, *388*, 594.
[59] Brange, J.In *The Physico-chemical and Pharmaceutical Aspects of Insulin*, Springer- Verlag, Berlin, **1987**, pp. 17.
[60] Krayenbuhl, C.; Rosenberg, T. *Reports of the Steno Memorial Hospital and the Nordisk Insulin Laboratories*, **1946**, *1*, 60.
[61] Coffman, F.D.; Dunn, M.F. *Biochemistry,***1988**, *27*, 6179.
[62] Roy, M.; Brader, M.L.; Lee, R.W.K.; Kaarsholm, W.C.; Hansen, J.F.; Dunn, M.F. *J. Biol. Chem.,* **1989**, *264*, 19081.
[63] Hansen, J.F. *Biophys. Chem.,* **1991**, *39*, 107.
[64] Wollmer, A.; Rannefield, B.; Stahl, J.; Melberg, S.G. *Hoppe-Seyler's Z Biol. Chem.,* **1989**, *370*, 1045.
[65] Bentley, G.A.; Brange, J.; Derewenda, Z.; Dodson, E.J.; Dodson, G.G.; Markussen, J.; Wilkinson, A.J.; Wollmer, A.; Xiao, B. *J. Mol. Biol.,***1992**, *228*, 1163.
[66] Hallden, G.; Gafvellin, G.; Mutt, V.; Jornvall, H. *Arch. Biochem. Biophys.,* **1986**, *247*, 20.
[67] Zimmerman, A.E.; Yip, C.C. *J. Biol. Chem.,* **1974**, *249*, 4021.
[68] Marki, F.; de Gasparo, M.; Eisler, K.; Kamber, B.; Riniker, B.; Rittel, W.; Sieber, P. *Hoppe-seyler's Z. Physiol. Chem.,* **1979**, *360*, 1619.
[69] Adams, M.J.; Dodson, G.; Dodson, E.; Hodgkin, D.C. In *Conformation of Biopolymers;* Ramachandran, G.N. Ed.; Academic Press, New York, **1967**, Vol. *1*, pp 9.
[70] Emdin, S.O.; Dodson, G.; Cutfield, J.M.; Cutfield, S.M. *Diabetologia,* **1980**, *19*, 174.
[71] Sudmeier, J.L.; Bell, S.J.; Storm, M.C.; Dunn, M.F. *Science,* **1981**, *212*, 560.
[72] Storm, M.C.; Dunn, M.F. *Biochemistry,* **1985**, *24*, 1749.
[73] Alameda, G.K.; Evelhoch, J.L.; Sudmeier, J.L.; Birge, R.R. *Biochemistry,* **1985,** *24*, 1757.
[74] Dunn, M.F.; Palmieri, R.; Kaarlsholm, N.C.; Roy, M.; Lee, R.; Dauter, Z.; Hill, C.; Dodson, G.G. In *Proceedings of the 5th International Symposium on Calcium Binding Proteins in Health and Disease*. Norman, A.W., Vanaman, T.C.; Means, A.; Eds.; Academic, New York, **1987**, 3872-3883.
[75] Palmieri, R.; Lew, R.W.K.; Dunn, M.F. *Biochemistry,* **1988**, *27*, 3387.
[76] Hill, C.P.; Dauter, Z.; Dodson, E.J.; Dodson, G.G.; Dunn, M.F. *Biochemistry,* **1991**, *30*, 917.
[77] Herman, L.; Sato, T.; Hales, C.H. *J. Ultrastruct. Res.,* **1973**, *42*, 298.
[78] Xiao, B.; Dodson, G.G.; Dodson, E.J.; Clarkson, J.; Turkenburg, J.; Derewenda, U.; Brady, L.; Brange, J. *J. Cryst. Growth,* **1992**, *122*, 144.
[79] Smith, P.K. *Molecular Genetics*. Macmillan Press. Hong Kong, **1991**.