Binding of Insulin to Its Receptor: Towards an Understanding in Three Dimensions

Ursula Schell,[a] Manfred Grün,[b] and Rolf Hilgenfeld*[a]

KEYWORDS:
electron microscopy · insulin · receptors · signal transduction · structure elucidation

More than 142 million people in the world suffer from diabetes mellitus; according to the World Health Organisation, this number is expected to double by the year 2025. This chronic disease, a dysfunction of carbohydrate and lipid metabolism, is caused by the failure of the pancreas to produce (sufficient amounts of) insulin. Diabetes can also be due to mutations in the insulin receptor which prevent the hormone from binding to and signalling through it. Thus, both sufficient availability of insulin and a functioning insulin receptor are essential for glucose homeostasis.

While external application of insulin has been used for the treatment of diabetes patients since the early 1920s, it has so far not been possible to match the insulin levels in healthy individuals as they vary throughout the day, depending on glucose uptake. Until now, the focus of the development of insulin analogues has been on generating variants with a modified ability to self-associate. A prominent example is LysPro insulin[1a] which shows a significantly reduced self-association and fast absorption upon subcutaneous injection. A long-acting insulin, Hoe901, will be introduced into the market this year. Strong interhexameric contacts introduced by "crystal contact engineering"[1b] are responsible for stabilization of Hoe901 crystals formed in the subcutaneous space, thereby generating an insulin depot with protracted glucose-lowering activity, and resulting in a much improved basal insulin supply.[1c,d]

While these new developments certainly contribute to an increased quality-of-life, several injections of fast-acting insulins before meals plus one injection of a long-acting preparation will continue to be the day-to-day reality for most type I diabetes patients. At present, an orally applicable insulin preparation remains a dream. However, chances are that it could become reality one day, if we continue to improve our understanding of the interactions between insulin and its receptor, thereby facilitating the design of small-molecule insulin agonists. Today, there is no full consensus as to what the receptor binding sites of the hormone are (see below), and even the 1:1 stoichiometry of the interaction is debated by some. This is due to the severe lack of structural information on the insulin – receptor complex. Efforts to crystallize this complex for X-ray structural analysis have been made throughout the past two decades, with little success. It became increasingly clear that a "divide and conquer" approach would be more practical. This route was recently followed by Luo et al.,[2] who determined the tertiary and quaternary structure of the insulin – receptor complex at low resolution by cryoelectron microscopy and interpreted their findings with the help of the available X-ray and NMR structures of isolated domains.

The insulin receptor is a disulfide-linked $(\alpha\beta)_2$ dimer of molecular mass 480 kDa. Each monomer consists of an α- and a β-chain, again linked by a disulfide bond.[3] The α-chain is extracellular, whereas the β-chain consists of a short extracellular portion (which, together with the entire α-chain, constitutes the ectodomain of the receptor), a single transmembrane segment, and an intracellular tyrosine kinase domain. Thus, the insulin receptor differs from most other tyrosine kinase receptors by being in an intrinsic dimeric state even in the absence of the ligand. Binding of insulin to the extracellular α-chain results in autophosphorylation of specific tyrosines in the cytoplasmic domain and the initiation of an intracellular signal transduction cascade.[4] However, the structural basis for insulin receptor activation upon ligand binding is a riddle yet to be solved.

Thanks to the efforts of X-ray crystallographers and NMR spectroscopists, three-dimensional structures for some building blocks of the insulin receptor are available today. In 1998, the crystal structure of the first three N-terminal domains of the α-chain of the homologous receptor for insulin-like growth factor 1 (IGF1) was solved.[5a] This fragment, which comprises part of the IGF1 binding site (although it fails to bind the hormone), consists of the domains L1, cysteine-rich, and L2 (together known as the "LCL region"). The same domains are present in the insulin receptor α-chain.[3a] The rest of the ectodomain largely comprises three fibronectin type III domains, to which both the α- and β-chains contribute. Structures of several homologous domains in other proteins, including fibronectin itself, have been determined by X-ray crystallography and NMR spectroscopy.[5b–d] Finally, the crystal structures of the intracellular tyrosine kinase domain with and without a bound ATP analogue and a peptide substrate have been determined,[5e,f] providing a high-resolution picture of the conformational changes likely to occur in this domain upon receptor activation.

Previous electron microscopic (EM) studies have been carried out on deter-

[a] *Prof. Dr. R. Hilgenfeld, Dr. U. Schell*
Department of Structural Biology & Crystallography
Institute of Molecular Biotechnology
Beutenbergstrasse 11, 07745 Jena (Germany)
Fax: (+49) 3641-65-6062
E-mail: hilgenfd@imb-jena.de

[b] *Prof. Dr. M. Grün*
JenaDrugDiscovery GmbH
Löbstedter Strasse 78, 07749 Jena (Germany)

  1439-4227/00/01/01 $ 17.50+.50/0

gent-released whole receptors[6a] and solubilized receptors reconstituted into lipid.[6b] The images of the whole receptor suggested a thin T-shaped molecule.[6a,b] In contrast, images reported for the soluble ectodomain alone showed variable shapes.[6c,d] Some light was recently shed onto the ectodomain structure by the preparation of Fab – ectodomain complexes which revealed a U-shaped form, containing a cleft between the two arms wide enough to accommodate a ligand.[6e] These data also provided the first experimental evidence for the internal symmetry model proposed by De Meyts.[7] Compared to previous studies, the approach of Luo et al.[2] offers a number of advantages. In contrast to negative staining, cryofixation, as applied in their work, should prevent any severe chemical modification and the receptor molecules can be expected to be in a more functional state. Application of STEM (scanning transmission electron microscopy) with a beam size of only 3 Å at $-150\,^{\circ}C$ and a low dose of 6 e Å^{-2} yields a higher resolution (20 Å) and low sample damage. For identifying and delimiting the insulin-binding site, a 70-atom gold cluster was coupled to the N terminus of the B chain (residue Phe-B1) of bovine insulin. This residue is not directly involved in receptor binding.[8a] The 3D reconstruction of the insulin – receptor complex at the full expected volume is—in contrast to previous studies—compact and globular, while at an intermediate density threshold, single domains become visible with a strong twofold vertical symmetry (Figure 1). The high density of the gold label identifies the extracellular region in the reconstruction and for most of the complexes, it reveals the binding of just one insulin molecule per receptor molecule.

What do we know about the insulin binding site? Can the electron microscopic 3D reconstruction help us see conformational changes upon insulin binding? It has been shown previously by use of a photoreactive insulin that the latter binds to the LCL region of the insulin receptor.[8b] These three domains (see above) form the minimal binding fragment of the receptor, and are responsible for the initial, low-affinity ligand-binding event.[8c] Surprisingly enough, it was shown by the crystal structure of the LCL region of the IGF1 receptor[5a] that the L domains exhibit a parallel β-helix architecture with α-helices capping the N- and C-terminal ends (Figure 2). The cysteine-rich domain consists of a series of disulfide-linked modules. A large cleft is formed by the three domains, which is believed to be the site of ligand binding. Facing the cleft is a β-sheet from the underside of L1 and an extended loop from the cysteine-rich domain. The size of the cleft as observed in the structure is slightly too large to bind IGF1 tightly; yet an approximately 25° rotation of the L2 domain towards the L1 domain would be sufficient to produce a cleft in which all three domains could contact IGF1.[5a] A rotation of the L2 domain was suggested to be one element of the signal transduction mechanism used by the IGF1 and insulin receptors. Unfortunately, the reso-

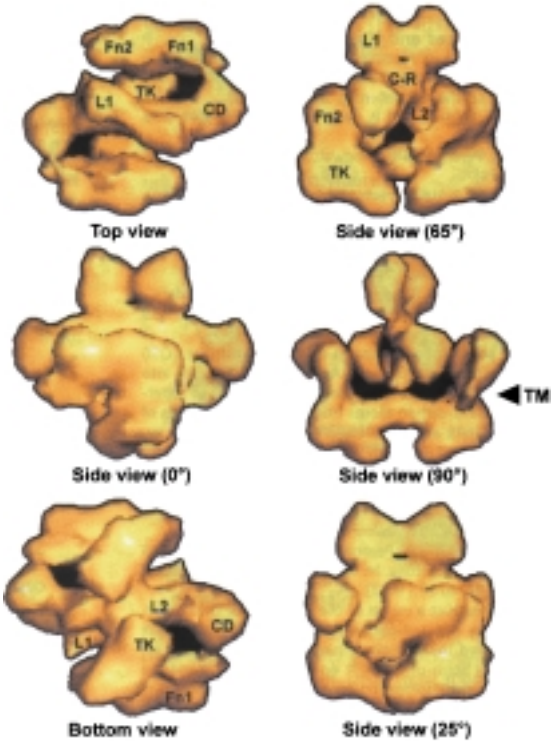

Figure 1. 3D Reconstruction of the insulin – receptor complex with twofold symmetry, shown at approximately 70% of full volume. Labels, for only one $\alpha\beta$ monomer, refer to biochemical domains. The arrowhead (TM) indicates the proposed position of the cell membrane. L1, C-R, and L2 = L1, cysteine-rich, and L2 domains, respectively; CD = "connecting domain" (a putative fibronectin III domain); Fn1 and Fn2 = fibronectin III domains 1 and 2; TK = tyrosine kinase; TM = transmembrane domain. (Reprinted with permission from Luo et al., Science ***1999***, *285, 1077 – 1080. ©American Association for the Advancement of Science, 1999).*

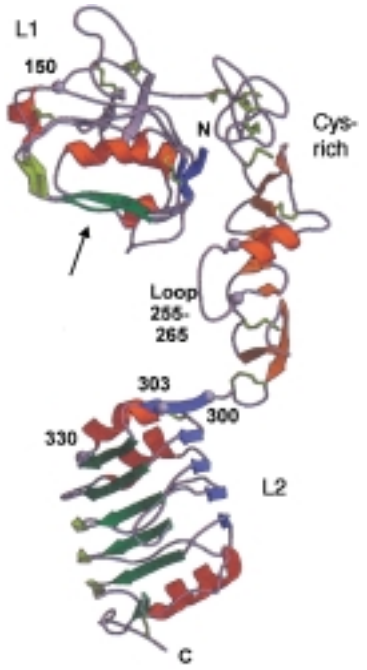

Figure 2. Ribbon diagram of the L1, cysteine-rich, and L2 domains of the insulin-like growth factor 1 (IGF1) receptor showing the putative ligand binding site between the three domains. Residues important for ligand binding are located on the underside of the L1 domain (indicated by the arrow) and in the loop of the cysteine-rich domain. (Reprinted with permission from Garrett et al., Nature ***1998***, *394, 395 – 399. ©Nature Macmillan Publishers Ltd., 1998).*

lution of the electron microscopic 3D reconstruction by Luo et al.[2] (Figure 3) is not yet high enough to make such movement visible.

Why is there only one insulin monomer bound to the dimeric insulin receptor? Luo et al.[2] observe a slightly asymmetric location of the gold-labeled insulin on the receptor surface and suggest insulin binding to the L1 and cysteine-rich domains of one α-subunit and to the L2 domain of the other (Figure 3b). This is also supported by previous results reviewed by McInnes and Sykes.[8d] Investigations with many site-specific mutants have generated information on the residues and regions of insulin involved in receptor binding, although it is not easy to make a distinction between each residue's role in receptor interaction and in ensuring the integrity of the fold of the hormone. Current belief is that there is a primary receptor binding site comprising residues A1 – A5, A16, A19, and, perhaps, A21 in the insulin A chain, and B12, B15, B16, B23, B24, B25 and, possibly, B26 in the B chain.[8e–g] A secondary site consisting of leucines A13 and B17 has also been described.[8h] Obviously, two distinct binding surfaces on insulin are required to achieve cross-linking to the binding sites on each of the α-subunits of the receptor.[8h,i] This would also explain the negative cooperativity of insulin binding[8j] and the low binding affinity of insulin receptor monomers.[8k,l]

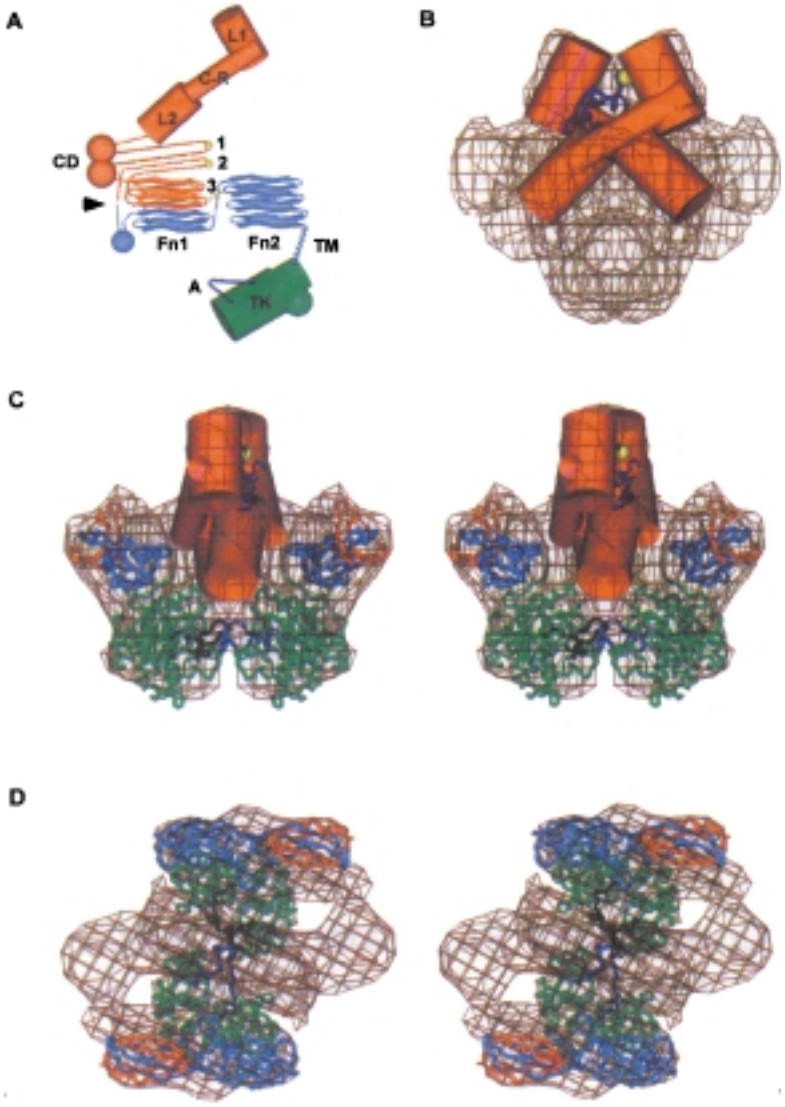

Figure 3. *Fitting of the known X-ray structures of domains to the 3D reconstruction from electron microscopy. a) Schematic presentation of the domain organization for one $\alpha\beta$ monomer. Structures not to scale; α-subunit, red; β-subunit, blue and green; unknown structures, spheres or lines. Labels: A = activation loop, 1, 2, and 3 = positions of specific cysteine residues. Other labels as described in Figure 1. b) Fitting of LCL domains as approximate cylinders to the insulin receptor ectodomain (wire mesh representation). One insulin molecule (purple ribbon) is shown inserted with its receptor-binding sites contacting the L1 and cysteine-rich domains (red) of one α-subunit and the L2 domain of the other. The gold marker (yellow) coincides with a high-density site. c) Right-angle view of (b) with LCL domains, fitted tyrosine kinase structure (green),[5e,f] two dimeric FnIII structures (blue and red), A loop (black) of the left tyrosine kinase domain in crystallographic position, and A-loop (dark-blue) of symmetry-related kinase extended to overlap peptide substrate of opposite kinase.[10] d) Right-angle top view of (c) without the LCL domains. One wire mesh square is 6.5 Å. (Reprinted with permission from Luo et al.,* Science ***1999***, 285, *1077 – 1080. ©American Association for the Advancement of Science, 1999).*

Prior to the work of Luo et al.,[2] the relative orientation of the fibronectin type III (FnIII) domains in the extracellular juxtamembrane region was totally unknown. There are three such domains in tandem, one each in the α- and β-subunits and one shared between the two.[9a,b,c] However, Luo et al.[2] did not observe them in a linear array as in the structure of fibronectin itself,[5b] but in a much more compact form (Figure 3a). These domains form at least two interchain disulfide bonds[9d] and may play a role in dimerization.[9e] The L2-proximal FnIII domain, frequently also called the "connecting domain", was suggested to be involved in the signal transmission to the tyrosine kinase.[9f]

What does the combined information from electron microscopy and X-ray crystallography tell us about the cytoplasmic domain? The overall architecture of the tyrosine kinase domain comprises an N-terminal lobe whose predominant structural feature is a five-stranded β-sheet, and a larger C-terminal lobe which is mainly α-helical. ATP binds in the cleft between the two lobes, and the tyrosine-containing substrate binds to the C-terminal lobe. Three tyrosine autophosphorylation sites are present in the activation loop of the domain (Tyr 1162/1163/1158). In the absence of insulin, a conformation of this loop is favored with Tyr 1162 engaged in the active site and both substrate and ATP binding sites being inaccessible.[5e] Upon ligand binding, however, a conformational change places the activation loop within reach of the active site of the other subunit in the dimer, thus permitting *trans* autophos-

phorylation.[10] The latter results in a dramatically different loop conformation (Figure 4).[5f] This allows unrestricted access to the binding sites for ATP and protein substrates (neighboring insulin receptor β-chain and downstream substrates such as IRS-2), and facilitates the proper spatial arrangement of residues involved in MgATP coordination. The electron microscopic study of Luo et al.[2] is not in disagreement with these ideas but lacks the resolution to provide much support either.

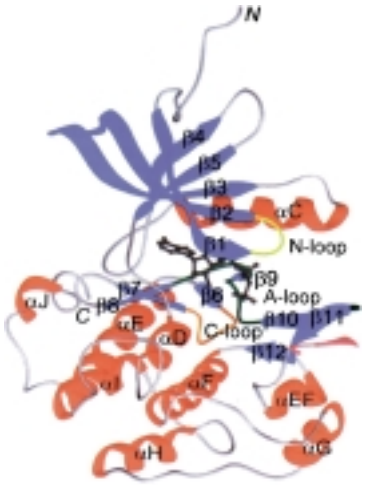

Figure 4. *Ribbon diagram of the phosphorylated tyrosine kinase domain of the insulin receptor. The α-helices are shown in red, the β-strands in blue, the nucleotide binding loop (N-loop) in yellow, the catalytic loop (C-loop) in orange, the activation loop (A-loop) in green, the ATP analogue AppNHp in black, and the peptide substrate in pink (ribbon below β11). The termini are denoted by N and C. (Reprinted with permission from Hubbard,* EMBO J. ***1997****,* 16*, 5572 – 5581. ©Oxford University Press, 1997).*

We now have an almost complete three-dimensional picture of the mutual arrangement of domains that compose the insulin receptor. However, further improvement of resolution of the model will be required to finally shed some light on the process of signal transmission. A first step in that direction has been achieved by the work discussed here.[2] The newly gained knowledge of the tertiary and quaternary structures may assist crystallization studies and help us to obtain X-ray grade crystals of the whole ectodomain or even the whole receptor, and of the complexes with insulin, thus making real high-resolution studies possible. On the other hand, one has to be aware of the limitations of cryoelectron microscopy, not only in terms of resolution. For example, the fixation of functionally intact receptor molecules on a support may not necessarily yield a realistic model. Thus, Luo et al.[2] indicate that the membrane-spanning portion of their 3D reconstruction is much thicker than one would expect.

In summary, it can be safely stated that the known pieces of the puzzle fit nicely together, but that several important parts are still missing. Further progress in understanding the structural basis of receptor activation, including more accurate determination of the molecular recognition sites on both the ligand and the receptor, will in the future allow us to focus on the development of insulin analogues with a modified receptor interaction. New and exciting advancements in the design of therapeutic agents for the treatment of diabetes almost certainly lie ahead.

[1] a) R. D. DiMarchi, R. E. Chance, H. B. Long, J. E. Shields, L. J. Slieker, *Horm. Res.* **1994**, *41*, S93-S96; b) R. Hilgenfeld, H. Berchtold, unpublished results; c) M. Dreyer, M. Pein, C. Schmidt, B. Heidtmann, M. Schlünzen, R. Roßkamp, *Diabetologia* **1994**, *37 (Suppl. 1)*, A78; d) H. Berchtold, R. Hilgenfeld, *Biopolymers* **1999**, *51*, 165 – 172.

[2] R. Z.-T. Luo, D. R. Beniac, A. Fernandes, C. C. Yip, F. P. Ottensmeyer, *Science* **1999**, *285*, 1077 – 1080.

[3] a) M. Bajaj, M. D. Waterfield, J. Schlessinger, W. R. Taylor, T. Blundell, *Biochim. Biophys. Acta* **1987**, *916*, 220 – 226; b) A. Ullrich, J. Schlessinger, *Cell* **1990**, *61*, 203 – 212.

[4] M. F. White, C. R. Kahn, *J. Biol. Chem.* **1994**, *269*, 1 – 4.

[5] a) T. P. J. Garrett, N. M. McKern, L. Meizhen, M. J. Frenkel, J. D. Bentley, G. O. Lovrecz, F. C. Elleman, L. J. Cosgrove, C. W. Ward, *Nature* **1998**, *394*, 395 – 399; b) D. J. Leahy, I. Aukhil, H. P. Erickson, *Cell* **1996**, *84*, 155 – 164; c) J. M. de Pereda, G. Wiche, R. C. Liddington, *EMBO J.* **1999**, *18*, 4087 – 4095; d) V. Copie, Y. Tomita, S. K. Akiyama, S. Aota, K. M. Yamada, R. M. Venable, R. W. Pastor, S. Krueger, D. A. Torchia, *J. Mol. Biol.* **1998**, *277*, 663 – 682; e) S. R. Hubbard, L. Wei, L. Ellis, W. A. Hendrickson, *Nature* **1994**, *372*, 746 – 754; f) S. R. Hubbard, *EMBO J.* **1997**, *16*, 5572 – 5581.

[6] a) K. Christiansen, J. Tranum-Jensen, J. Carlsen, J. Vinten, *Proc. Natl. Acad. Sci. USA* **1991**, *88*, 249 – 252; b) J. Tranum-Jensen, K. Christiansen, J. Carlsen, G. Brentex, J. Vinten, *J. Membr. Biol.* **1994**, *140*, 215 – 223; c) J. D. Johnson, M. L. Wong, W. J. Rutter, *Proc. Natl. Acad. Sci. USA* **1988**, *85*, 7516 – 7520; d) E. M. Schaefer, H. P. Erickson, M. Federwisch, A. Wollmer, L. Ellis, *J. Biol. Chem.* **1992**, *267*, 23 393 – 23 402; e) P. A. Tulloch, L. J. Lawrence, N. M. Mc Kern, C. P. Robinson, J. D. Bentley, L. Cosgrove, N. Ivancic, G. O. Lovrecz, K. Siddle, C. W. Ward, *J. Struct. Biol.* **1999**, *125*, 11 – 18.

[7] P. De Meyts, *Diabetologia* **1994**, *37*, 135 – 148.

[8] a) J. Murray-Rust, A. N. McLeod, T. L. Blundell, S. P. Wood, *BioEssay* **1992**, *14*, 325 – 331; b) M. Fabry, E. Schaefer, L. Ellis, E. Kojro, F. Fahrenholz, D. Brandenburg, *J. Biol. Chem.* **1992**, *267*, 8950 – 8956; c) C. Kristensen, F. C. Wiberg, L. Schaffer, A. S. Andersen, *J. Biol. Chem.* **1998**, *273*, 17 780 – 17 786; d) C. McInnes, B. D. Sykes, *Biopolymers* **1997**, *43*, 339 – 366; e) E. N. Baker, T. L. Blundell, J. F. Cutfield, S. M. Cutfield, E. J. Dodson, G. G. Dodson, D. H. C. Hodgkin, R. E. Hubbard, N. W. Isaacs, C. D. Reynolds, K. Sakabe, N. Sakabe, N. M. Vijayan, *Philos. Trans. R. Soc. London Ser. B* **1988**, *319*, 369 – 456; f) S. Gammeltoft, *Physiol. Rev.* **1984**, *64*, 1321 – 1378; g) C. Kristensen, T. Kjieldsen, F. C. Wiberg, L. Schaffer, M. Hach, S. Havelund, J. Bass, D. F. Steiner, A. S. Andersen, *J. Biol. Chem.* **1997**, *272*, 12 978 – 12 983; h) L. Schaffer, *Eur. J. Biochem.* **1994**, *221*, 1127 – 1132; i) M. F. White, *Diabetologia* **1997**, *40*, 2 – 17; j) M. Böni-Schnetzler, W. Scott, S. M. Waugh, E. DiBella, P. F. Pilch, *J. Biol. Chem.* **1987**, *262*, 8395 – 8401; k) F. Canals, *Biochemistry* **1992**, *31*, 4493 – 4501; l) J. Schlessinger, A. Ullrich, *Neuron* **1992**, *9*, 383 – 391.

[9] a) C. Marino-Buslje, K. Mizuguchi, K. Siddle, T. L. Blundell, *FEBS Lett.* **1998**, *441*, 331 – 336; b) T. D. Mulhern, G. W. Booker, L. Cosgrove, *Trends Biochem. Sci.* **1998**, *23*, 465 – 466; c) C. W. Ward, *Growth Factors* **1999**, *16*, 315 – 322; d) L. G. Sparrow, N. M. McKern, J. J. Gorman, P. M. Strike, C. P. Robinson, J. D. Bentley, C. W. Ward, *J. Biol. Chem.* **1997**, *272*, 29 460 – 29 467; e) P. Bork, K. Downing, B. Kieffer, I. D. Campbell, *Q. Rev. Biophys.* **1996**, *29*, 119 – 167; f) C. K. Sung, K. Y. Wong, C. C. Yip, D. M. Hawley, I. D. Goldfine, *Mol. Endocrinol.* **1994**, *8*, 315 – 324.

[10] S. R. Hubbard, *Prog. Biophys. Mol. Biol.* **1999**, *71*, 343 – 358.