Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Crystallization and preliminary crystallographic investigation of a low-pH native insulin monomer with flexible behaviour

Youshang Zhang,[a] Jean L. Whittingham,[b] Johan P. Turkenburg,[b] Eleanor J. Dodson,[b] Jens Brange[c] and G. Guy Dodson[b]*

[a]Shanghai Institute of Biochemistry and Cell Biology, Academia Sinica, 320 Yue-Yang Road, Shanghai 200031, People's Republic of China, [b]York Structural Biology Laboratory, Department of Chemistry, University of York, Heslington, York YO10 5DD, England, and [c]Brange Consult, Krøyersvej 22C, 2930 Klampenborg, Denmark

Correspondence e-mail: ggd@ysbl.york.ac.uk

Received 6 October 2001
Accepted 31 October 2001

Insulin naturally aggregates as dimers and hexamers, whose structures have been extensively analysed by X-ray crystallography. Structural determination of the physiologically relevant insulin monomer, however, is an unusual challenge owing to the difficulty in finding solution conditions in which the concentration of insulin is high enough for crystallization yet the molecule remains monomeric. By utilizing solution conditions known to inhibit insulin assembly, namely 20% acetic acid, crystals of insulin in the monomeric state have been obtained. The crystals are strongly diffracting and a data set extending to 1.6 Å has recently been collected. The crystals nominally belong to the space group *I*422, with unit-cell parameters $a = b = 57.80$, $c = 54.61$ Å, giving rise to one molecule in the asymmetric unit. Preliminary electron-density maps show that whilst most of the insulin monomer is well ordered and similar in conformation to other insulin structures, parts of the *B*-chain C-terminus main chain adopt more than one conformation.

1. Introduction

Since the isolation of insulin in 1922, this peptide hormone has been the subject of extensive clinical, biochemical and biophysical studies. Human insulin, which consists of an *A* and a *B* chain of 21 and 30 amino-acid residues, respectively, interacts with the receptor as a monomer, yet one if its most striking characteristics is its ability to form different association states including dimers, tetramers and hexamers. There appears to be a good physiological explanation for this: hexamer formation helps to ensure complete conversion of proinsulin to insulin, facilitates efficient packing of insulin in the pancreatic storage vesicles and may protect the protein from physical and chemical degradation (Dodson & Steiner, 1998; Brange, 1994).

Structural analyses on insulin show how the monomer is adapted for dimer and hexamer formation (Baker *et al.*, 1988). The general consistency of the monomer structure in dimers, hexamers and other assemblies seen in the earlier studies gave the misleading impression that the protein conformation observed in these structures largely represents the receptor-binding conformation. It was only when the crystal structure of a completely inactive insulin, made so by introducing a peptide cross-link between the *A*-chain N-terminus and the *B*-chain C-terminus, turned out to have exactly the same conformation as that of those obtained in earlier structural studies that it was realised that the *B*-chain C-terminus undergoes a considerable conformation change on binding to the receptor (Derewenda *et al.*, 1991). In contrast, it has been demonstrated that the N-terminus of the *A* chain needs to retain its observed α-helical structure for the insulin molecule to remain active (Weiss *et al.*, 2000).

Insulin dimer formation is driven by the burial of hydrophobic residues on the molecule and by the favourable β-strand interactions that are generated between the two *B*-chain C-termini in the dimer. Consequently, native insulin can only be obtained as a monomer (the state in which the hormone binds to its receptor) at less than micromolar concentrations, which is too low for most biophysical studies. Therefore, to study and crystallize native insulin as a monomer has been very difficult. Various tactics have been employed to prevent dimerization. One method, which has proved very successful, is to introduce mutations at the dimer-forming surface which interfere with dimerization without having a detrimental effect on activity. For example, the mutations Tyr*B*16→His or Pro*B*28→Asp/Lys have produced stable monomeric insulins suitable for solution studies including NMR (Ludvigsen *et al.*, 1994). In addition, it has been known for some time that insulin dimerization can be totally prevented by dissolving native insulin in 20% acetic acid, even at high concentrations such as 20 mg ml^{-1}. These conditions have made possible some revealing studies on the insulin monomer in solution which show that the molecule is very mobile, with flexible surface elements (Weiss *et al.*, 1989). Since this behaviour generally dis-

favours crystal formation, no attempt was made to use the acetic acid conditions for the crystallization of insulin. However, we have recently obtained insulin crystals in the presence of acetic acid, resulting in the first crystal structure of a native (porcine) insulin monomer.

2. Experimental procedure

2.1. Crystallization and data collection

Crystals were grown by the hanging-drop vapour-diffusion method. The hanging drops consisted of 2 µl of an insulin solution [20 mg ml^{-1} in 20%(*v*/*v*) acetic acid] mixed with 2 µl of a reservoir solution [20%(*v*/*v*) acetic acid and 0.05 *M* sodium sulfate]. Each drop was suspended over 1 ml of the reservoir solution. Crystals were obtained at 291 K after 2–3 d (Fig. 1). A single crystal of dimensions 0.2 × 0.2 × 0.1 mm was flash-frozen at 120 K in a cryoprotectant solution consisting of 60% reservoir and 40% ethylene glycol prior to data collection on station ID14-3 at the ESRF, Grenoble. A 1.6 Å data set was collected in a single sweep of 75°, using a 1.0° oscillation angle per image. The data were processed using *XDISP*, *DENZO* and *SCALEPACK* from the *HKL* package (Otwinowski & Minor, 1997) and *SCALEPACK2MTZ*, *TRUNCATE* and *CAD* from the *CCP*4 program suite (Collaborative Computational Project, Number 4, 1994).

2.2. Preliminary structure determination

The monomeric insulin structure was solved by molecular replacement using *AMoRe* (Navaza, 1994) as implemented in the *CCP*4 suite (Collaborative Computational Project, Number 4, 1994). The monomeric subunit of the 2-zinc pig insulin hexamer (Baker *et al.*, 1988) was used as the search model, resulting in a final correlation factor after fitting of 0.50 and a crystallographic *R* factor of 0.51. Five cycles of maximum-likelihood refinement were then carried out using *REFMAC* (Collaborative Computational Project, Number 4, 1994), after which the crystallographic *R* factor had dropped to 0.42 ($R_{\rm free}$ = 0.46). $2F_o - F_c$ and $F_o - F_c$ electron-density maps were then examined using *XFIT* in *QUANTA* (Oldfield, 2001) and preliminary rebuilding of the structure was carried out.

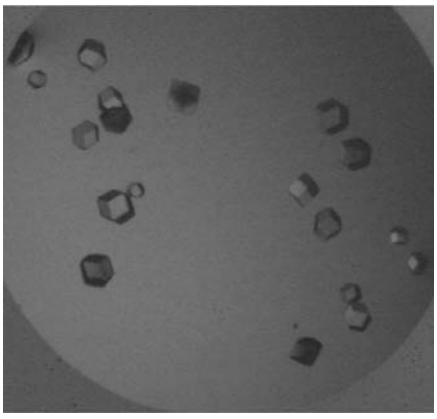

Figure 1
Crystals of porcine insulin grown in 20% acetic acid and 0.05 *M* sodium sulfate.

3. Results and discussion

The data were indexed in space group *I*422, with unit-cell parameters $a = b = 57.80$, $c = 54.61$ Å. The crystals contain one monomer in the asymmetric unit, corresponding to a solvent content of 37.0%. The data set was 98.7% complete between 20.0 and 1.6 Å resolution, with an overall merging *R* value of 0.03. Other data statistics are shown in Table 1. The crystal packing interactions of the newly solved structure revealed that the insulin monomer, in the absence of normal dimer formation, was able to form an alternative packing arrangement supported by interactions between the *B*9–*B*19 α-helices of two adjacent monomers centred around residue His*B*10. After some preliminary refinement, $2F_o - F_c$ and $F_o - F_c$ electron-density maps showed that the *A*-chain residues and the *B*-chain residues from Asn*B*3 to Gly*B*20 were well defined. However, electron density for the *B*-chain C-terminal residues *B*25–*B*30 was essentially absent, indicating that this part of the molecule is highly flexible when not involved in dimer formation. This observation supports the notion that the *B*-chain C-terminus undergoes a considerable conformational rearrangement during receptor binding.

The insulin monomers are arranged in the crystal in such a way that the *B*-chain C-termini are in close proximity around the crystallographic fourfold axis, making well ordered packing of residues *B*26–*B*30 sterically impossible. For residues *B*21–*B*25 there was fragmented electron density indicative of multiple conformations for this part of the molecule. However, subsequent refinement of the overall structure did not improve the electron-density maps around residues *B*20–*B*25 sufficiently to allow model building in this region. It was then considered that the close proximity of these residues to the crystallographic fourfold axis might coincide with a breakdown in the *I*422 symmetry of the molecule in this region and that reassignment of the crystal to a lower symmetry space group might allow individual conformations at *B*20–*B*25 to be identified. Therefore, it will be necessary to recollect the data and resolve the structure in a lower symmetry space group. The resulting structure will be described in detail in a future paper.

Table 1
Data-processing statistics for monomeric insulin.

Resolution limits	No. of unique reflections	$\lvert I/\sigma(I)\rvert$	Completeness (%)	Multiplicity	$R_{\rm sym}$
20.00–4.33	363	53.3	98.4	5.0	0.025
4.33–3.44	334	64.4	99.4	5.6	0.017
3.44–3.01	334	53.6	99.4	5.6	0.019
3.01–2.73	309	53.5	100.0	5.8	0.023
2.73–2.54	333	49.9	98.8	5.8	0.029
2.54–2.39	322	40.0	99.4	5.8	0.036
2.39–2.27	315	37.6	98.7	5.9	0.040
2.27–2.17	314	33.5	99.1	5.9	0.051
2.17–2.09	311	33.8	98.7	5.9	0.052
2.09–2.02	320	26.0	99.4	5.9	0.067
2.02–1.95	299	22.4	99.0	6.0	0.086
1.95–1.90	308	18.7	98.4	6.0	0.106
1.90–1.85	325	15.8	98.5	6.0	0.123
1.85–1.80	297	13.6	97.4	6.0	0.142
1.80–1.76	309	12.1	99.4	5.9	0.164
1.76–1.72	317	9.8	98.1	6.1	0.196
1.72–1.69	306	8.3	97.1	6.0	0.226
1.69–1.66	300	7.2	99.3	6.1	0.265
1.66–1.63	307	5.0	96.5	6.0	0.387
1.63–1.60	299	4.8	98.4	5.9	0.379
All *hkl*	6322	47.2	98.7	5.9	0.030

We would like to thank the BBSRC and Novo Nordisk A/S, Denmark for providing the financial support and insulin material for these studies. We would like to acknowledge the use of station ID14-3 at the ESRF and thank Lisa Wright and Koen Verschueren for collecting the data. We are also very grateful to Anthony Wilkinson, Gideon Davies and Michael Weiss (Case Western Reserve University, Cleveland) for useful and enjoyable discussions.

References

Baker, E. N., Blundell, T. L., Cutfield, J. F., Cutfield, S. M., Dodson, E. J., Dodson, G. G., Crowfoot Hodgkin, D. M., Hubbard, R. E., Isaacs, N. W., Reynolds, C. D., Sakabe, K., Sakabe, N. & Vijayan, N. M. (1988). *Philos. Trans. R. Soc. London*, **319**, 369–456.
Brange, J. (1994). *Stability of Insulin.* Dordrecht, The Netherlands: Kluwer Academic Publishers.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Derewenda, U., Derewenda, Z., Dodson, E. J., Dodson, G. G., Xiao, B. & Markussen, J. (1991). *J. Mol. Biol.* **220**, 425–433.
Dodson, G. G. & Steiner, D. (1998). *Curr. Opin. Struct. Biol.* **8**, 189–194.
Ludvigsen, S., Roy, M., Thøgersen, H. & Kaarsholm, N. C. (1994). *Biochemistry*, **33**, 7998–8006.
Navaza, J. (1994). *Acta Cryst.* A**50**, 157–163.
Oldfield, T. J. (2001). *Acta Cryst.* D**57**, 82–94.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Weiss, M. A., Hua, Q. X., Jia, W. H., Chu, Y. C., Wang, R. Y. & Katsoyannis, P. G. (2000). *Biochemistry*, **39**, 15429–15440.
Weiss, M. A., Nguyen, D. T., Khait, I., Inouye, K., Frank, B. H., Beckage, M., O'Shea, E., Shoelson, S. E., Karplus, M. & Neuringer, L. J. (1989). *Biochemistry*, **28**, 9855–9873.