*Macedonian Journal of Chemistry and Chemical Engineering*, Vol. **34**, No. 1, pp. 67–72 (2015)

MJCCA9 – 664

ISSN 1857-5552
e-ISSN 1857-5625

Received: Januray 5, 2015
Accepted: February 23, 2015

UDC: 548.55:548.73
*Original scientific paper*

# CRYSTAL STRUCTURE OF THE COBALT HUMAN INSULIN DERIVATIVE♦

**Biserka Prugovečki*, Natalija Ivetić, Dubravka Matković-Čalogović***

*University of Zagreb, Faculty of Science, Department of Chemistry, Division of General and Inorganic Chemistry, Horvatovac 102a, 10000 Zagreb, Croatia*

*biserka@chem.pmf.hr, dubravka@chem.pmf.hr

The structure of the human cobalt insulin derivative at 1.73 Å resolution is described. Single crystals were prepared by the hanging drop vapour diffusion crystallization method using Zn-free insulin and cobalt(II) acetate.

The crystal structure was determined by the single crystal X-ray diffraction method. The investigated insulin derivative exhibits the $T_6$ form of insulin and crystallizes in the trigonal system in space group *R*3, with the unit cell parameters $a = b = 81.43$ Å and $c = 33.75$ Å. There are two cobalt atoms per insulin hexamer which are octahedrally coordinated by three symmetry-related $N^{\epsilon 2}$ atoms of three HisB10/HisD10 and three oxygen atoms from three water molecules.

**Keywords***:* insulin derivative; cobalt; X-ray structure

## КРИСТАЛНА СТРУКТУРА НА ХУМАН КОБАЛТЕН ИНСУЛИНСКИ ДЕРИВАТ

Решена е кристалната структура на хуман кобалтен инсулински дериват со резолуција од 1,73 Å. Монокристалите се добиени со помош на методот на парна дифузија по принципот на висечка капка, користејќи инсулин (без Zn) и кобалт(II)ацетат. Кристалната структура е решена со рендгенска дифракција на монокристал. Изучуваниот дериват на инсулинот претставува форма $T_6$ на инсулин и кристализира во тригоналниот систем, просторна група *R*3, со параметри на елементарната ќелија, $a = b = 81{,}43$ Å и $c = 33{,}75$ Å. Во инсулинскиот хексамер постојат два кобалтни атома кои се октаедарски координирани со три симетриски поврзани атоми $N^{\epsilon 2}$ од три HisB10/HisD10 и три кислородни атоми од три молекули на вода.

**Клучни зборови:** инсулински дериват; кобалт; рендгенска структура

## 1. INTRODUCTION

Insulin is a hormone that regulates the carbohydrate metabolism and takes part in the metabolism of fat and proteins. It is used medically in patients with Type 1 diabetes mellitus and occasionally in some patients with Type 2 diabetes mellitus. Due to its crucial metabolic role and its pharmaceutical importance many structural studies on chemically and genetically modified insulins have been done [1]. Insulin was discovered by Banting and Best in 1921 and has since been involved in many landmarks in the development of biology [2]. For instance, insulin was the first protein to be sequenced [3]. Insulin is structured as two polypeptide chains, chain A consisting of 21 and chain B of 30 amino acids. Chain A and chain B are linked by two disulphide bridges, while the third intra chain disulphide bridge links residues A6 and A11. Porcine insulin differs from human insulin by only one amino acid (at the position B30 there is Ala in porcine and Thr in human insulin) and bovine insulin is also very similar, differing in only three positions (AlaA8, ValA10, AlaB30 in

♦ Dedicated to Academician Gligor Jovanovski on the occasion of his 70th birthday.

bovine and ThrA8, IleA10, ThrB30 in human). Because of their similarity, these forms of insulin are also recognized by our own cells and may be used in therapy, although human insulin is now prevailing. Insulin is accumulated in the pancreas as a $Zn^{2+}$ containing hexamer and insulin hexamers are also used in therapeutic preparations for the control of diabetes, whereas the monomer is the active form. There are three forms of insulin hexamers, $T_6$, $T_3R_3$ and $R_6$. Kaarshom proposed the use of the terms T (tense) and R (relaxed), by analogy to the situation in haemoglobin, to describe the conformation of the structures rather than the zinc content in insulin [4]. In $T_6$ insulin all six monomers in the hexamer have B1 to B8 extended, while in the $T_3R_3$ insulin hexamer B1 to B8 chain in three out of six molecules becomes helical, giving rise to the $T_3R_3$ state. Monoclinic and rombohedral phenol insulin crystals can be obtained in the $R_6$ state in which all six monomers have B1 to B8 helical.

$T_6$ hexamers are formed in the absence of high chloride ion concentrations or phenolic derivatives. In the $T_6$ hexamer insulin molecules there are two $Zn^{2+}$ ions lying on a three-fold axis, each of them octahedrally coordinated by three symmetry dependent nitrogen atoms from the histidine side chain HisB10 (or HisD10) and by three water molecules [5, 6].

It has been shown that the chloride ion concentration of 6% or more induces the T to R transition at residues B1-B8 of the insulin molecule. In crystals of native insulin grown with high chloride ion concentrations, three more off-axial sites for zinc binding occur, with B5His residues taking part in the coordination of the zinc [7]. The off-axial sites are not always fully occupied and it was found that in the $T_3R_3$ type of hexamer there can be a different number of Zn ions per hexamer with different coordination [8, 9].

In the native $R_6$ hexamer, zinc ions are only found on the three-fold axis, tetrahedrally coordinated by three symmetry dependent nitrogen atoms from the histidine side chain His B10 and usually an anion. 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 [10, 11].

It has been known for many years that different metal ions can substitute the zinc ions in the 2Zn hexamer [12], however only a few such complexes have been structurally characterized [13–18].

As a part of our ongoing research on structural characterization of insulin derivatives, in the present study the $Zn^{2+}$ ions in human insulin were substituted with $Co^{2+}$ ions in order to investigate the difference in metal coordination and conformation of the insulin molecule.

## 2. EXPERIMENTAL

### 2.1. *Crystallization*

Crystals of the cobalt derivative were grown by the hanging-drop method originally established by Cutfield [19] and improved by Xiao [20]. Biosynthetic human Zn-free insulin was supplied by Lilly Research Laboratories. Salts and other reagents were purchased and used without further purification. Optimum crystallization conditions were as follows: the protein solution consisted of 7.5 mg $ml^{-1}$ Zn-free insulin in 0.02 M HCl, while the reservoir solution contained 1 mM solution of sodium citrate, pH 6.4, $\varphi$ (acetone) = 10%, 16.5 mM solution of cobalt(II)acetate and redistilled water. Each drop consisted of 1 μl of protein solution and 1 μl of reservoir solution. It took about 7 days for crystals to grow to the final size. This insulin derivative crystallizes in the rhombohedral space group *R*3, with cell dimensions $a$ = 81.43, $c$ = 33.75 Å.

### 2.2. *Diffraction data collection*

Intensity data for cobalt insulin derivative were collected from the single crystal of dimensions 0.31 × 0.25 × 0.08 $mm^3$ (Fig. 1).

**Fig. 1**. Crystals of the Co human insulin derivative (left); the single crystal used for data collection (right)

The single crystal was cryoprotected by immersion in solution consisting of 70% (v/v) reservoir solution and 30% (v/v) glycerol. Diffraction data were collected on a laboratory diffractometer Rigaku RU-H3R rotating-anode generator equipped with Osmic multilayer optics (wavelength, λ = 1.5418 Å), and with the Mar 345 detector from MarResearch and Oxford Cryosystems cryocooler.

A 1.73 Å data set was collected, with oscillation angles of 1.0°. The data were processed with MOSFLM [21] and then scaled and merged by use of the CCP4 suite [22]. Data statistics are given in Table 1.

T a b l e 1

*Data measurement and refinement statistics for the $T_6$ human cobalt insulin derivative*

| **Data measurement** | |
|---|---|
| Space group | *R*3 |
| Unit cell parameters | |
| $a=b$, Å | 81.43 |
| $c$, Å | 33.75 |
| $\alpha=\beta$ / ° | 90 |
| $\gamma$ / ° | 120 |
| Temperature, K | 100 |
| Resolution range, Å | 20.00-1.73 |
| No. of unique reflections (all) | 26503 |
| *R* free test set | 453 refl. (5.2%) |
| Overall $R_{merge}$ | 0.057 |
| $R_{merge}$ in high resolution shell (1.82–1.73 Å) | 0.230 |
| Mean ($I/\sigma I$) | 9.7 (3.0 in the high resolution shell) |
| **Refinement statistics**[a] | |
| Resolution range, Å | 20.00–1.73 (1.77–1.73) |
| Number of reflections (observed) | 8210 (565) |
| Completeness | 99.46 (94.29) |
| *R* | 0.146 (0.197) |
| $R_{free}$ | 0.206 (0.333) |
| Total number of atoms | 998 |
| Average *B* (all atoms), Å$^2$ | 18.0 |
| Wilson B-factor, Å | 14.0 |
| Bulk solvent *k*sol/eÅ$^{-3}$, *B*sol, Å$^2$ | 0.45, 57.9 |
| R.m.s. deviations from ideal | |
| Bond lengths, Å | 0.018 |
| Bond angles, ° | 1.797 |
| Dihedral angles, ° | 6.998 |

[a] Values in parentheses are for the highest resolution shell.

### 2.3. *Structure determination and refinement*

The coordinates of human Zn insulin (PDB entry 1mso) from which all water molecules, zinc ions and alternate side chains were omitted, was taken as a starting model. Refinement was carried out using the maximum-likelihood minimization implemented in REFMAC [23], with 5% of the total data being excluded from the refinement and used for calculating $R_{free}$. The structure was then refined to a higher resolution and throughout the refinement $2F_o–F_c$ and $F_o–F_c$ maps were calculated and examined by the COOT program [24] where manual adjustments were made. A total of 111 water molecules were added by the ARP/wARP program [25] in combination with examination in COOT. Only the two cobalt atoms were refined anisotropically. Low electron density was observed for the terminal atoms of the side chain of PheD1 and ValD2, however they could be modelled as disordered in two positions. This region is usually disordered and sometimes cannot be modelled at all, as in the structure of porcine Co-insulin where these two amino acids were not included in the model.

Some other side chains were also found to be disordered, and alternate conformations were added for AsnA21, GlnB4, ValB12, GluB21, AsnC21, ValD18, LysD29 and ThrD30.

## 3. RESULTS AND DISCUSSION

The only structure with cobalt that has replaced the zinc ion is that of the porcine Co-insulin derivative published by Nicholson *et al.* [14] (PDBID 1M5A). They have prepared this derivative by replacement of zinc with cobalt using a 1.5 molar excess of cobalt(II) acetate in the crystallization process, whereas we have used Zn-free human insulin and so had no zinc present in the crystallization solution.

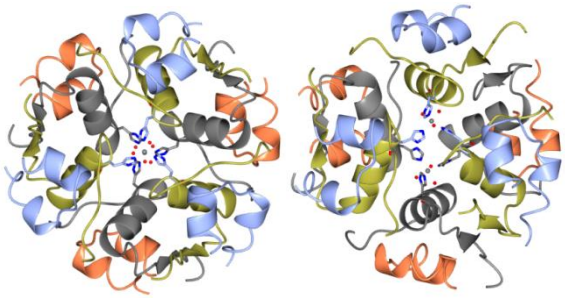

**Fig. 2**.Cartoon drawing of the human Co-insulin derivative. Chain A is shown in blue, B in green, C in orange and D in grey. Left: the view is down the three-fold axis and two $Co^{2+}$ ions are one above the other (grey sphere); right: side view with both $Co^{2+}$ ions

### 3. 1. *Coordination of the cobalt ion*

The crystal structure revealed that the hexamer is in the $T_6$ form with two $Co^{2+}$ ions lying on the three-fold axis so it can also be referred to as the 2Co-insulin (Fig. 2). The cobalt ion is octahedrally coordinated by three symmetry dependent nitrogen atoms from the histidine side chain His B10 (or HisD10) and by three water molecules (Fig. 3). This coordination sphere is the same as in the native 2Zn-insulin and other metal derivatives such as manganese [18] and nickel [15] that were crystallized in low chloride concentration (less than 6%). The structure of the porcine 2Co-insulin

revealed the same coordination sphere of the cobalt ion with the Co-N distances of 2.08 and 2.16 Å. In the present structure the bond lengths Co-N are 2.15 and 2.21 Å and the bond lengths to the water molecules are 2.09 and 2.43 Å. The bond lengths Zn-N in the Zn-insulin derivative are 2.09 and 2.10 Å while the Zn-O bond lengths are 2.20 and 2.23 Å. Only three small-molecule crystal structures with $Co^{2+}$ ions coordinated by three histidine ligands and three water molecules have been deposited in the *Cambridge Structural Database* [26] (BOVMIJ, LONQEL and TUCBAV) refs. [27–29]. In these structures the Co-N distance is in the range from 2.074(2) to 2.166(2) Å which is comparable to the same distance in the protein structures.

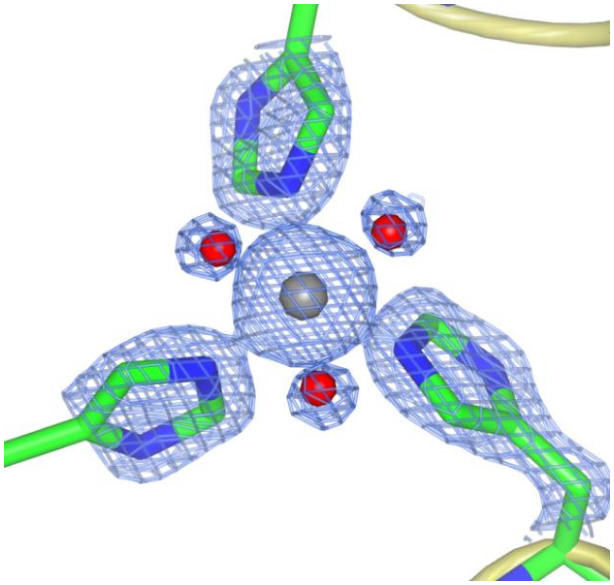

**Fig. 3**. Octahedral coordination of the cobalt atom ($Co^{2+}$ion is shown in grey colour) with the electron density map (2Fo–Fc map; contour level 1.08 eÅ$^{-3}$, shown as a blue net). The view is down the three-fold axis.

### 3.2. *Hexamer conformation and differences with the porcine Co-derivative and human Zn-derivative*

The monomers form a hydrogen bonded antiparallel β-sheet in the dimer with hydrogen bond contacts of the type N-H···O between insulin monomers at PheB24···TyrD26, PheD24···TyrB26, TyrB26···PheD24 and TyrD26···PheB24 (Table 2, Fig. 4). The asymmetric unit consists of one $T_2$ dimer and by the three-fold symmetry axis the $T_6$ hexamer is assembled. The root mean square differences (r.m.s.d.) were calculated by the LSQKAB program [30] in the CCP4 program suite. Unexpectedly, a greater difference was found between the human and porcine cobalt derivatives than the human cobalt and zinc derivatives. The greatest r.m.s.d. difference of 0.436 Å between the main chain atoms was found between the B chains of the human and porcine Co-insulin derivatives (Table 3). The greatest difference within the two human structures of the Co- and Zn-derivatives, both in the main chain and within all atoms is in the D chain.

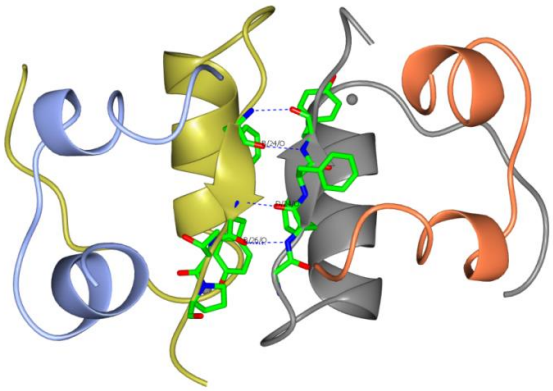

**Fig. 4**. Hydrogen bonds between two monomers that give a dimer

T a b l e 2

*Intermolecular hydrogen bonds that connect monomers into a dimer*

| Neighb. (N1) | Hbond atom (B1) | Hbond atom (B2) | Neighb (N2) | Angle(N1-B1-B2)/° | d(B1···B2)/Å | Angle (B1-B2-N2)/° |
|---|---|---|---|---|---|---|
| B/26(TYR)/CA | B/26(TYR)/N | D/24(PHE)/O | D/24(PHE)/C | 112.68 | 2.87 | 164.48 |
| D/26(TYR)/CA | D/26(TYR)/N | B/24(PHE)/O | B/24(PHE)/C | 118.85 | 2.85 | 171.07 |
| B/26(TYR)/C | B/26(TYR)/O | D/24(PHE)/N | D/24(PHE)/CA | 142.71 | 3.01 | 136.59 |
| D/26(TYR)/C | D/26(TYR)/O | B/24(PHE)/N | B/24(PHE)/CA | 154.83 | 2.87 | 134.90 |

Nicholson *et al.* argued that their structure (the porcine Co-derivative) is indicative of an early stage of conformational change required in the T to R transition. They observed large shifts in dihedral angles for the residues B5 to B9. However, superposition of the main chains of three $T_6$ structures (the present cobalt derivatives, human, 4RXW, and porcine, 1M5A, and the human zinc derivative, 1MSO [6]) with one $T_3R_3$ structure (human Zn-derivative; 1G7A [31]) shows that all three $T_6$ struc-

tures are very similar whereas the $T_3R_3$ structure is significantly different in the B chain and especially in the D chain (Fig. 5). Any intermediate structure that would imply this transition cannot be seen.

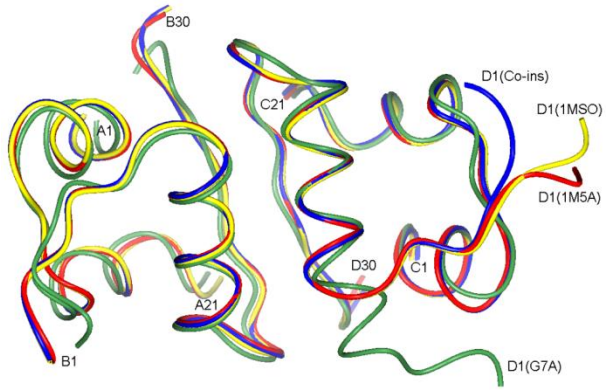


**Fig. 5**. Superposition of the A, B, C and D chains in four insulin structures of which three are of the $T_2$ type, 4RXW (human Co-insulin derivative – present structure; blue), 1M5A (porcine Co-insulin derivative; red), 1MSO (human Zn-insulin; yellow), and one is of the TR type, 1G7A (human Zn-insulin; green)

Table 3

*Root mean square differences (Å) between the investigated $T_6$ human Co-insulin derivative (4RXW) and Zn-insulin (1MSO) or porcine Co-insulin (1M5A)*

| | Co-insulin (1M5A) | | Zn-insulin (1MSO) | |
|---|---|---|---|---|
| Residue range | All atoms | Main chain | All atoms | Main chain |
| A chain | 0.681 | 0.167 | 0.295 | 0.124 |
| B chain | 0.824 | 0.436 | 0.534 | 0.249 |
| C chain | 0.492 | 0.192 | 0.417 | 0.145 |
| D chain | 0.800 | 0.432 | 0.922 | 0.365 |

## 4. CONCLUSION

The results show that the investigated human cobalt insulin derivative adopts the $T_6$ conformation with two cobalt ions per hexamer located on the 3-fold axis. Both cobalt ions are octahedrally coordinated by three symmetry-related $N^{\varepsilon 2}$ atoms of three HisB10/HisD10 and three oxygen atoms from three water molecules. The structure is similar to the $T_6$ Zn-insulin derivative and also to the porcine Co-insulin derivative, as was expected, with the largest root mean square difference for the B and D chains of the porcine Co-insulin derivative of 0.436 and 0.432, respectively, for the main chain atoms. The highest root mean square difference is smaller in the case of the human Zn-insulin derivative being 0.365 for the D chain.

***Supplementary material.*** The structure was deposited at the *Protein Data Bank*, with PDB reference code 4RXW.

***Acknowledgement.*** This work was supported by the Ministry of Science, Education and Sports of the Republic of Croatia grant No. 119-1193079-1084.

## REFERENCES

[1] H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, P. E. Bourne, The Protein Data Bank, *Nucl. Acids Res.*, **28**, 235–242 (2000).www.rcsb.org

[2] F. G. Banting, C. H. Best, The internal secretion of the pancreas, *J. Lab. Clin. Med* .**7**, 251–266 (1922).

[3] F. Sanger, E. O. P. Thompson, The amino-acid sequence in the glycyl chain of insulin, *Biochem. J.* **53**, 353–366 (1953).

[4] N. C. Kaarsholm, H. C. Ko, M. F. Dunn, Comparison of solution structural flexibility and zinc binding domains for insulin, proinsulin, and miniproinsulin, *Biochemistry* **28**, 4427–4435 (1989).

[5] M. J. Adams, T. L. Blundell, E. J. Dodson, G. G. Dodson, M. Vijayan, E. N. Baker, M. M. Harding, D. C. Hodgkin, B. Rimmer, S. Sheat, Structure of rhombohedral 2 zinc insulin crystals, *Nature* **224**, 491–495 (1969).

[6] G. D. Smith, W. A. Pangborn, R. H. Blessing, The structure of $T_6$ human insulin at 1.0 Å resolution, *Acta Crystallogr.* **D59**, 474–482 (2003).

[7] G. Bentley, E. Dodson, G. Dodson, D. Hodgkin, D. Mercola, Structure of insulin in 4-zinc insulin, *Nature* **261**, 166–168 (1976).

[8] G. D. Smith, D. C. Swenson, E. J. Dodson, G. G. Dodson, C. D. Reynolds, Structural stability in the 4-zinc human insulin hexamer, *Proc. Nat. Acad. Sci. USA*, **81**, 7093–7097 (1984).

[9] E. Ciszak, G. D. Smith, Crystallographic evidence for dual coordination around zinc in the $T_3R_3$ human insulin hexamer, *Biochemistry* **33**, 1512–1517 (1994).

[10] J. W. Whittingham, S. Chaudhuri, E. J. Dodson, P. C. E. Moody, G. G. Dodson, X-ray crystallographic studies on hexameric insulins in the presence of helix-stabilizing agents, thiocyanate, methylparaben and phenol, *Biochemistry* **34**, 15553–15563 (1995).

[11] G. D. Smith, E. Ciszak, L. A. Magrum, W. A. Pangborn, R. H. Blessing, $R_6$ hexameric insulin complexed with *m*-cresol or resorcinol, *Acta Crystallogr.* **D56**, 1541–1548 (2000).

[12] C. P. Hill, Z. Dauter, E. J. Dodson, G. G. Dodson, M. F. Dunn, X-ray structure of an unusual $Ca^{2+}$ site and the roles of $Zn^{2+}$ and $Ca^{2+}$ in the assembly, stability, and storage of the insulin hexamer, *Biochemistry* **30**, 917–924 (1991).

[13] B. Prugovečki, Coordination of selected essential metal ions in insulin, PhD Thesis. (In Croatian, Abstract in English), University of Zagreb, 2005.

[14] J. Nicholson, L. Perkins, F. Körber, The high resolution structure of a hexameric $T_6$ cobalt insulin: A possible pathway for the T to R transition, *Recent Res. Dev. Mol. Biol.* **3**, 1–16 (2006).

[15] B. Prugovečki, E. J. Dodson, G. G. Dodson, D. Matković-Čalogović, Structure of the $T_6$ human nickel insulin derivative at 1.35 Å resolution, *Croat. Chem. Acta* **82(2),** 433–438 (2009).

[16] R. Sreekanth, V. Pattabhi, S. S. Rajan, Metal induced structural changes observed in hexameric insulin, *Int. J. Biol. Macromol* .**44**, 29–36 (2009).

[17] N. R. S. Krishna, V. Pattabhi, S. S. Rajan, Metal induced conformational changes in human insulin: Crystal structures of $Sr^{2+}$, $Ni^{2+}$ and $Cu^{2+}$ complexes of human insulin, *Protein & Peptide Letters* **18**, 457–466 (2011).

[18] B. Prugovečki, I. Pulić, M. Toth, D. Matković-Čalogović, High resolution structure of the manganese derivative of insulin, *Croat. Chem. Acta* **85(4)**, 435–439 (2012).

[19] S. M. Cutfield, D. Phil. Thesis, University of Oxford, 1975.

[20] B. Xiao, D.Phil. Thesis, University of York, 1990.

[21] A. G. W. Leslie, Recent changes to the MOSFLM package for processing film and image plate data, Joint CCP4 + ESF-EAMCB Newsletter on Protein Crystallography, No. 26, 1992, pp. 27–33.

[22] M. D. Winn, C. C. Ballard, K. D. Cowtan, E. J. Dodson, P. Emsley, P. R. Evans, R. M. Keegan, E. B. Krissinel, A. G. W. Leslie, A. McCoy, S. J. McNicholas, G. N. Murshudov, N. S. Pannu, E. A. Potterton, H. R. Powell, R. J. Read, A. Vagin, K. S. Wilson, Overview of the *CCP*4 suite and current developments, *Acta. Crystallogr.*, **D67**, 235–242 (2011).

[23] G. N. Murshudov, A. A.Vagin, E. J. Dodson, Refinement of macromolecular structures by the maximum-likelihood method, *Acta Crystallogr.*, **D53**, 240–255 (1997).

[24] P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, Features and development of *Coot*,*Acta Crystallogr.*, **D66**, 486–501 (2010).

[25] G. Langer, S. X. Cohen, V. S. Lamzin, A. Perrakis, Automated macromolecular model building for X-ray crystallography using *ARP/ wARP* version 7. *Nature Protocols.*, **3**, 1171–1179 (2008).

[26] F. H. Allen, The Cambridge Structural Database: A quarter of a million crystal structures and rising, *Acta Crystallogr.*, **B58**,380–388 (2002).

[27] J.-J. Nie, J.-H. Li, D.-J.Xu, Hexakis(1*H*-imidazole-$KN^3$)cobalt(II)triaquatris(1H-imidazole-$KN^3$)cobalt(II) bis(naphthalene-1,4-dicarboxylate), *Acta Crystallogr.*, **E65**, m822–m823 (2009).

[28] P. Liu, Z. Liu, Triaqua-tris(1*H*-imidazole)-cobalt(II) 2,6-naphthalenedisulfonate monohydrate, *Z. Kristallogr. NCS*, **223,** 498–500 (2008).

[29] T. K. Yazıcılar, E. Y. Gürkan, I. Uçar, C. Kazak, Synthesis, characterization, spectroscopic, and thermal studies of imidazole complexes of metal benzenesulfonates. Crystal structure of $[M(imH)_3(H_2O)_3]\cdot(BS)_2$ [M=Ni(II) and Co(II), *Transition Met. Chem.* **34,** 669–676 (2009).

[30] W. Kabsch, *Acta Crystallogr.*, A solution for the best rotation to relate two sets of vectors, **A32,** 922–92 (1976).

[31] G. D. Smith, W. A. Pangborn, R. H. Blessing, Phase changes in $T_3R_3^f$ human insulin: temperature or pressure induced? *Acta Crystallogr*. **D57**, 1091–1100 (2001).