Supramolecular Chemistry
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713649759

Understanding the Peculiarities of Azide and Thiocyanate Binding in Proteins: Use of the Small Molecule Structural Data

Luba Tchertanov[a]
[a] Institut de Chimie des Substances Naturelles, Gif-sur-Yvette, FRANCE

To cite this Article Tchertanov, Luba(2000) 'Understanding the Peculiarities of Azide and Thiocyanate Binding in Proteins: Use of the Small Molecule Structural Data', Supramolecular Chemistry, 12: 1, 67 — 91
To link to this Article: DOI: 10.1080/10610270008029805
URL: http://dx.doi.org/10.1080/10610270008029805

SUPRAMOLECULAR CHEMISTRY, Vol. 12, pp. 67-91
Reprints available directly from the publisher
Photocopying permitted by license only

Understanding the Peculiarities of Azide and Thiocyanate Binding in Proteins: Use of the Small Molecule Structural Data

LUBA TCHERTANOV*

Institut de Chimie des Substances Naturelles, UPR 2301, CNRS, 91198 Gif-sur-Yvette, FRANCE

The complexation and supramolecular recognition of two polyfunctional anions, thiocyanate and azide, have been analysed using the X-ray structural data retrieved from the Brookhaven Protein Database (for macromolecules) and the Cambridge Structural Database (for small molecule structures). The following structural aspects have been considered: – hydrogen-bond acceptor function of the anions; – analysis of the coordination features of anions with the metals; – influence of metal complexation on the H-donor accepting properties of the anions. Lastly, a comparison of the modes of binding of the two anionic substrates was carried out. Their extraordinary activity as hydrogen-bond acceptors allows these anions to serve as a strong bridge between different molecular fragments. The coordination of thiocyanate or azide anion by the metal atoms in coordination compounds gives rise to a redistribution of anionic charge in both entities which in turn, controls the hydrogen-bonding acceptor properties of the terminal atom of the anion. The results presented in this paper provide structural information of the role of anion binding in proteins and to our knowledge, afford the first study of the binding behaviour of thiocyanate and azide in systematic manner.

Keywords: Thiocyanate, Azide, Metal Complexation, H-bonding, Enzyme, Statistical Analysis of Structural Data

INTRODUCTION

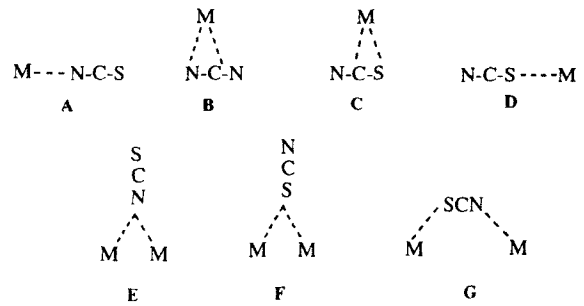
In the chemistry of life processes, anion binding plays a central role since it concerns essential aspects such as the activity of enzymes, the transport of hormones, protein synthesis and DNA regulation. There is, therefore, intensive effort being devoted to the problem of anion metal selectivity and non-covalent interaction phenomena (molecular recognition). Several comprehensive reviews on recent developments in anion complexation have been published [1]. Anions metal selectivity and non-covalent interaction phenomena have been investigated by numerous research groups [2]. The purpose of this paper is a study of metal complexation and molecular recognition of geometrically simple linear thiocyanate $(SCN)^-$ and azide N_3^- anions, which is extremely important in biochemistry. Indeed, both anions have been frequently used as crystallizing agents in protein crystal growth [3] and, much more importantly, are known to be competitive inhibitors in many enzymes. Understanding of the dynamics of biological

* Corresponding Author: Fax: 33-1-69-07-72-47, E-mail: Luba.Tchertanov@icsn.cnrs-gif.fr

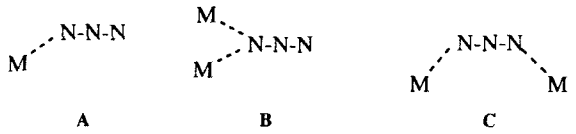
processes would entail thorough analysis of the geometrical features of anionic binding.

The examples of metalloproteins involving these small inorganic anions (e.g. superoxide dismutase, carbonic anhydrase etc.) clearly show that in each case, anion recognition is guided by attraction to the metal ion and/or hydrogen-bonding interactions with the donor-groups functionality. Preliminary analysis of anion sites in crystallographically determined macromolecules shows that in some cases the interpretation of the observed results is complicated or inadequate. It appears that the best way to study the interactions with anions is to analyze their environment through a survey of small-molecule crystal structures. Analysis of interactions between anion and metals and/or organic ligands using the Cambridge Structural Database [4] and the Brookhaven Protein Database [5] can provide structural evidence for the role of these anions.

Both entities, thiocyanate and azide, are small, triatomic, linear monoanions wherein the negative charge can be localized on one of the terminal atoms or delocalized over the whole anion. Both of them are the species with bidentate function and offer different modes of binding by metals and/or H-bond donor groups: through one of the terminal atoms (one or two times) or through both of them, by both ends, (bridging mode) [6]. The ambidentate character of thiocyanate (two different donor sites) and its bidentate function give rise to a wide variety of coordination geometries (Scheme 1).



SCHEME 1 Modes of metal-thiocyanate binding



SCHEME 2 Modes of metal-azide binding

In azide both terminal atoms are equivalent and its geometry is described only by two modes of binding: through one of the terminal nitrogen atoms (one or two times, **A** and **B**) or through both of them (bridging mode, **C**) (Scheme 2).

The following structural aspects have been studied: (i) peculiarities of coordination of anions by metal atoms (modes of coordination and selectivity); (ii) hydrogen-bond acceptor function of the anions and (iii) influence of metal complexation on the H-donor accepting properties of these anions. The first part of the paper gives a survey of the relevant protein structures which is mostly focused on the problems associated with establishing of anionic binding sites. The second part includes the detailed analysis of anion binding in small molecular structures. We compare data from all known structures to provide detailed structural information about the molecular recognition of azide and thiocyanate in macromolecules and small molecule structures. We also wished to compare the structural and accepting properties of azide and thiocyanate entities.

RESULTS

1. Anion-Binding Sites in Proteins

The analysis of the anion binding in proteins with the known crystal structures was performed. These negatively charged species can participate in a broad spectrum of non-covalent interactions, such as hydrogen bonding, van der Waals interactions and hydrophobic interac-

tions. What is actually happening around these triatomic anion in proteins? Are the anion further oriented by hydrogen bonding to neighboring atoms in the case of cation-anion binding? What specific properties are responsible for tight and selective binding of the ligands?

The PDB contains just a few examples of protein structures involving either thiocyanate (5 structures) or azide anions (14 structures); only half of them were reported in papers. All currently available protein structures containing these anions are listed in Tables I and II (for thiocyanate and azide respectively). These structures vary with respect to the functional roles of the proteins and the species from which they were derived; some of the proteins were chemically modified; others show several crystalline modifications. We will start with the analysis of the proteins containing the SCN units.

TABLE I Macromolecular structures containing thiocyanate entity

Refcode	*Protein*	*Fragment*	*Source*	*Resolution*	*Reference*
1TEW	*Lysozyme*	SCN	Turkey EGG white	1.65	[7]
6EBX	*Erabutoxin*-b	SCN	Sea-snake venome	2.0	[8]
2CA2	*Carbonic anhydrase II*	Zn-NCS	Human	1.9	[10]
2FAM	*Myoglobin*	Fe-SCN	Sea hare	2.0	[11]
2TCI	*Thiocyanate insulin*	Zn-NCS	Sus scrofa, pig	1.8	[9]

TABLE II Macromolecular structures containing azide entity

Code	*Protein*	*Fragment*	*Resolution*	*References*
1ASQ	*Ascorbate Oxidase* (Oxidoreductase)	NNN-Cu-NNN	2.3	[14]
1CVA	*Carbonic Anhydrase II* (Lyase, oxo-acid)	Zn-NNN	2.3	[17]
1HR3	*Hemerythrin* (Oxygen transport)	Fe-NNN	5.5	[18]
2HMZ	*Hemerythrin* (Adizomet) (Oxygen transport)	Fe-NNN	1.7	[18]
1ISC	*Iron(III)Superoxide Dismutase* (Oxidoreductase)	Fe-NNN	1.8	[12]
1MNG	*Mn Superoxide Dismutase* (Oxidoreductase)	Mn-NNN	1.8	[12]
1RAY	*Carbonic Anhydrase II* (Lyase, oxo-acid)	Zn-NNN	1.8	[17]
1SWM	*Mioglobin (Fe)* (Oxygen transport)	Fe-NNN	1.8	[19]
1UGB	*Human Carbonic Anhydrase II* (Lyase, oxo-acid)	Zn-NNN	2.0	[17]
1UGF	*Human Carbonic Anhydrase II* (Lyase(oxo-acid)	Zn-NNN	2.0	[17]
1VNC	*Chloroperoxidase* from the *Fungus* (Oxidoreductase)	V-NNN	2.1	[13]
2MHR	*Myohemerythrin* (Oxygen transport)	Fe-NNN	1.7	[18]
2NAD	*NAD-dependent formate dehydrogenase* (Oxidoreductase)	NNN	2.0	[16]
2TSB	*Azurin mutant M121A* (Electron transport)	Cu-NNN	2.3	[20]
5MBA	*Mioglobin (Fe)* (Oxygen storage)	Fe-NNN	1.9	[19]

1.1. *Thiocyanate binding*

Two proteins, lysozyme, **1TEW**, [7] and erabutoxin-b, **6EBX**, [8] were crystallised from thiocyanate and contain SCN in the anionic form. In the hexagonal form of turkey egg white lysozyme, the anion acts as a bridge between two symmetry related molecules: a pair of H-bonds involving the S atom (S...NH_2 and S...H_2O) links the anion with an arginine residue of the protein molecule and with a water molecule. The other terminal nitrogen atom, by means of a pair of H-bonds and through water molecules (N...HOH) binds the anion with a symmetry related molecule of protein. In erabutoxin-b, a post-synaptic neurotoxin isolated from the venom of the sea-snake *Laticauda semifasciata*, the thiocyanate anion is located in a polar cleft formed by three molecules (two crystallographically independent, A and B, and the third one symmetry related to A. The contacts N...OG Ser23B (2.68 Å) and N...NH Arg33A (3.26 Å) can be interpreted as H-bonds with polar residues. Sulfur is located as far as 3.31 Å from the oxygen of Cys54B and 3.43 Å from the water molecule which bridges two crystallographically independent molecules. Thus, thiocyanate is seen to act as an anionic bridge between the different protein molecules.

In two proteins, thiocyanate insulin, **2TCI**, [9] and carbonic anhydrase, **2CA2**, [10] we find SCN bonded to zinc. In hexameric insulin, the linear anion occupies a special position on the three-fold axis and is located between three symmetry related protein molecules. The bond distance Zn-N and bond angle ∠(ZnNC) are equal to 1.69 Å and 180° respectively. The terminal sulfur is located in a hydrophobic region. In the carbonic anhydrase complex, the SCN anion is bonded to the pentacoordinated zinc along with a water molecule and three histidine residues (Zn-N 1.91 Å, ∠(ZnNC) 173°). The terminal sulfur sits in a hydrophobic pocket, but is at a distance of 4.1 Å from a backbone NH (Thr199). This is the site of the "deep" water which is supposed to bind to Thr199 in the catalytic mechanism, wherein SCN acts as an inhibitor. We observe an intramolecular contact (HOH...N) between two neighbouring zinc ligands H_2O and NCS.

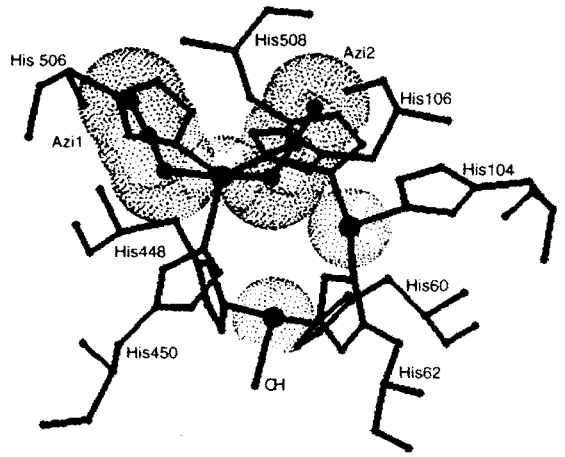


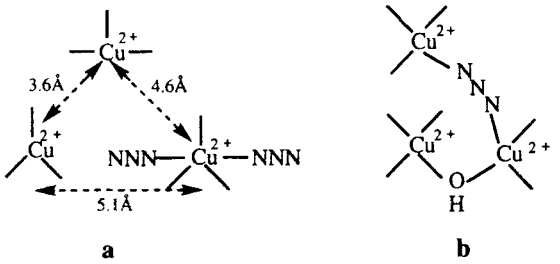
FIGURE 1 Trinuclear copper cluster in 1ASQ. Two azides (Azi1 and Azi2) coordinated to one copper cation as monodentate ligands (See Color Plate V at the back of this issue)

In the sea hare myoglobin (*Aplysia Limacina*) (ferric) complex with SCN, **2FAM** [11], SCN is coordinated to the *haem* iron through the S atom, and the terminal nitrogen forms two H-bonds with an arginine residue and a water molecule.

1.2. *Azide binding*

The enzymes with azide can be divided into several classes on the basis of their biological functions: oxidoreductases (**1ASQ, 1ISC, 1MNG, 1VNC** and **2NAD**); carbonic anhydrases (**1CVA, 1RAY, 1UGB** and **1UGF**); oxygen transport enzymes (**1HR3, 2HMZ, 1SWM, 5MBA**); electron transport enzyme (**2TSB**); oxygen binding enzyme(**2MHR**) (Table II).

Oxidoreductases: The manganese and the iron superoxide dismutases, **1ISC** from *escherichia coli* and **1MNG**, from *thermus thermophilus*, HB8 [12] have a very similar structure of metal-azide binding site. Both cations (Fe^{+3} and Mn^{+3}) are



SCHEME 3 Azide binding in the trinuclear copper cluster as derived from a) crystal structure of **1ASQ** ([14] and b) the spectroscopic model ([15]

six-coordinated by three histidines and one aspartate oxygen (carboxyl with monodentate mode) and by azide and a water molecule. Azide binds between two equatorial histidine ligands to complete a distorted octahedral coordination of metal centre. In the structure of haloperoxidase from the *fungus curvularia inaequalis* (**1VNC**) [13] a pentavalent vanadium cation is covalently bonded to the protein by histidine (His496) at one of the axial positions of the trigonal bipyramidal metal site. The azide binds to vanadium in the axial position, another axial and equatorial positions are occupied by hydroxides. Ascorbate oxidase from *zucchini* (**1ASQ**) [14] contains a trinuclear copper cluster in two crystallographically independent molecules. No bridges between the copper atoms in either of the clusters were observed, the Cu...Cu distances varying between 3.6 and 5.1 Å. Crystallographic data show that two azides are bound to the Cu atom through one terminal nitrogen (Figure 1 and Scheme 3, a). It should be noted, that the structure of trinuclear copper cluster with quite different (actually bridging) mode of azide binding was determined by magnetic circular dichroism study and reported in [15] (Scheme 3, b).

The asymmetric unit of **NAD-dependent formate dehydrogenase complex** (**2NAD** from *methylotrophic bacterium pseudomonas sp. 101*) [16] contains a dimer with two chemically identical subunits related by a non-crystallographic two-fold axis. NAD (nicotinamide-adenine-dinucleotide) is bound in the cleft separating the catalytic domains and mainly interacts with residues from the co-enzyme binding domain. An azide anion is located near the point of catalysis, in the proximity of NAD.

Carbonic anhydrases II (**1CVA, 1RAY, 1UGB** and **1UGF**) (*Homo Sapiens*) [17] catalyzes the reversible formation of bicarbonate and a proton from water molecule and carbon dioxide. Catalytic sites are mononuclear, the zinc cation is tetrahedral, coordinated by three His-N ligands (namely His94, His96 and His119) and azide anion. Although azide often facilitates the crystallization of carbonic anhydrase, this small anion does not cause any significant structural changes in the enzyme active site. Importantly, the azide binding at pH 8.0 has implication for the zinc-binding mode of the catalytic bicarbonate ion, as a competitive inhibitor.

Oxygen transport and binding proteins are structurally different. *Non-haem* iron proteins, as **hemerythrin** (**1HR3** from *siphonosoma species near siphonosoma funafati*; **2HMZ**, from *sipunculid worm*) and **myohemetytrin** (**2MHR** from *sipunculid worm)* [18] include the binuclear iron cluster which consists of two iron atoms bridged by a μ-oxygen atom. Both iron atoms are hexa-coordinated. Two carboxylate groups of a glutamate and an aspartate are bonded to both cations, the other ligands are nitrogen atoms of three histidines at one of the Fe cations, while the other one is coordinated by two hystidines and an azide anion. **Mioglobin** Fe (**1SWM** from *sperm whale* and **5MBA** from *sea hare*) [19] is a conjugated protein which is the oxygen transporting pigment of muscule. It is made up of one globin polypeptide chain and one *haem* group. This *haem* group contains the iron cation surrounded by four equatorial nitrogen atoms from protoporphyrin and two axial groups, histidine and azide. The azide is oriented towards the outer part of the distal site crevice.

Electron transport enzyme: Azurin mutant M121A (**2TSB** from *escherichia coli*) [20]. There are four molecules per asymmetric unit in this

form. Four copper clusters are quite similar. The coordination is best described as trigonal bipyramidal, with Cu^{+2} tetracoordinated by two hystidine ligands, azide and by sulfur of Cys112.

Evidently, the azide-containing proteins exhibit a magnificent variety of cation sites having a multitude of chemical and structural properties. Consequently, the azide roles differ both functionally and structurally. With the exception of NAD-dependent formate dehydrogenase complex (**2NAD**), in all extracted proteins azide is an exogenous ligand bonded to tetra-, penta- or hexacoordinated metal atom (Fe, Zn, Cu, Mn or V cations). In all macromolecules azide shows an unidentate mode of binding to metal (Scheme 2, **A**). Geometrical parameters of metal-azide interaction are listed in Table III. The values of cation-azide distance vary from 1.80 to 2.19 Å and bond angles MNN span the range from 90 to 145°.

TABLE III Geometry of azide binding sites in macromolecules. Metal-azide interaction

PDB code	*Cation*	*Unit*[a]	*Distance, (Å) M-N1*	*Angle, (°) MN1N2*	*Torsion, (°) MN1N2N3*
1ASQ	Cu^{+2}	**A: a**	2.19	124	146
	Cu^{+2}	**b**	1.93	108	−62
	Cu^{+2}	**B: a**	2.09	126	160
	Cu^{+2}	**b**	2.02	109	−114
1CVA	Zn^{+2}		2.04	92	175
2HMZ	Fe^{+3}	**A:**	2.12	123	100
	Fe^{+3}	**B:**	2.06	130	38
	Fe^{+3}	**C:**	2.14	125	−71
	Fe^{+3}	**D:**	2.04	124	−155
1ISC	Fe^{+3}	**A:**	2.09	117	−36
	Fe^{+3}	**B:**	2.16	116	−45
1MNG	Mn^{+3}	**A:**	2.30	141	−42
	Mn^{+3}	**B:**	2.16	145	−43
1RAY	Zn^{+2}		2.02	112	180
1SWM	Fe^{+2}		2.07	117	172
1UGB	Zn^{+2}		1.97	90	−177
1UGF	Zn^{+2}		1.80	116	−70
1VNC	V		1.98	136	165
2MHR	Fe^{+3}		2.11	127	−58
2TSB	Cu^{+2}	**A:**	2.00	142	131
	Cu^{+2}	**B:**	2.02	132	45
	Cu^{+2}	**C:**	2.07	144	177
5MBA	Fe^{+3}		2.08	122	176

a. **A, B**, etc are used to distinguish different subunits in the same protein molecule; **a** and **b** are used to distinguish different fragments in the same metal cluster.

We'll now focus our attention on other type of azide interactions, *viz.*, its non-covalent environment. It is to be emphasized that in most cases we do not have information on the hydrogen positions and we consider the appropriate groups as possible hydrogen donors if the distance N...HD(N/O) is within 3.4 Å (Table IV).

TABLE IV Geometry of azide binding sites in macromolecules. Non-covalent interactions around the fragment M-N1-N2-N3, (N...H-N/O ≤ 3.4 Å)

PDB code	*Unit*[a]	N_i	*Donor group*	*Distance, (Å) N_i...Donor*	*Angle, (°) DN1N2*
2NAD	**A:**	N1	Arg284(NH1)	3.12	163
		N1	Arg284(NH2)	3.38	125
		N1	His332(NE2)	3.34	119
		N3	Asn146(ND2)	2.79	139
		N3	Ile122(NH)	2.97	110
	B:	N1	Arg284(NH1)	2.99	159
		N1	Arg284(NH2)	2.99	120
		N1	His332(NE2)	3.34	119
		N3	Asn(ND2)	2.94	143
		N3	Ile122(NH)	2.94	120
1ASQ	**A: a**	N3	H_2O	2.53	121
		N3	H_2O	2.77	115
		N1	H_2O	2.98	108
	b	N3	H_2O	3.09	119
		N3	H_2O	3.34	143
	B: a	N3	H_2O	3.12	121
		N3	H_2O	3.30	136
	b	N1	H_2O	3.38	98
		N3	H_2O	2.51	113
1CVA		N3	Val199(NH)	3.16	103
1ISC	**A:**	N1	H_2O	3.08	110
		N3	His31(NH)	3.00	133
	B:	N3	His31(NH)	3.04	130
1MNG	**A:**	N3	Tyr36(OH)	1.60	101
		N3	H_2O	3.20	119
	B:	N3	Tyr36(OH)	2.60	115
		N3	H_2O	3.40	121
1RAY		N1	H_2O	2.55	124
		N3	Leu198(NH)	3.40	91
1SWM		N1	His64(NE2)	2.91	115
1UGF		N2	N_2O	2.69	101
		N3	Leu198(NH)	3.44	87
		N3	H_2O	3.26	115
1UGB		N1	Thr99(OG1)	2.67	100
5MBA		N3	Arg66(NH2)	2.79	146

a. **A, B,** etc are used to distinguish different subunits in the same protein molecule; **a** and **b** are used to distinguish different fragments in the same metal cluster.

The unique example of azide in discrete form is observed in the structure of **2NAD**. The structure of azide site in both crystallographically independent molecules is similar and only one molecule is presented in Figure 2, a. The azide anion binds with the protein ligands *via* multiple H-bonding involving both of its terminal nitrogen atoms (Figure 2, b and Table IV). One terminal nitrogen atom participates in the strong H-bonding with NH of main chain of Ile122 (N3... HN 2.97 in **A** and 2.94 Å in **B** molecules) and with amino group of Asn146 (N3... HN 2.79 in **A** and 2.94 Å in **B**). The second terminal nitrogen atom is in a very good position with Arg284. Both distances N1... HN (3.12 and 3.38 in **A** and 2.99 Å in **B**) may be interpreted as a three-centred hydrogen bonding. The distance N1..NE2 3.34 Å with histidine probably corresponds to a weak H-bonding. The azide unit is nearly parallel to the plane of the NAD nicotinamide aromatic cycle and the distances between azide atoms and the cyclic atoms are in the range of 2.80 – 4.30 Å. The short distance N3...H-C_{sp}^2 with Phe98 (N... HC 3.37 Å) is also noteworthy.

It is important to emphasize that in carbonic anhydrase II structures all zinc ligands paticipate in H-bonding. These non-covalent interactions increase the zinc affinity by a factor of about 10 [21]. In all structures azide is H-bonded differently: a) by the only terminal nitrogen (non-covalently bonded with a metal) with NH of main chain of Val199 (N3... HN of 3.16 Å) in **1CVA** (Scheme 4, a); b) by the only nitrogen atom bonded to Zn in H-bonding with Thr199 (N1...HO of 2.67 Å) in **1UGB** (Scheme 4, b); c) by both azide ends, with main chain of Leu189 (N3... HN of 3.40 Å) and a water molecule (N1...H_2O of 2.55 Å) (Scheme 4, c); d) by central azide nitrogen atom bonded to a water molecule (N2...H_2O of 2.69 Å) in **1RAY** and e) twice through one of the terminal atoms (N3...HO 3.36, N3... HN 3.40 Å Leu198) in **1UGF**.

In manganese and iron superoxide dismutases, **1ISC** and **1MNG**, azide is H-bonded mainly througth the terminal nitrogen atom.

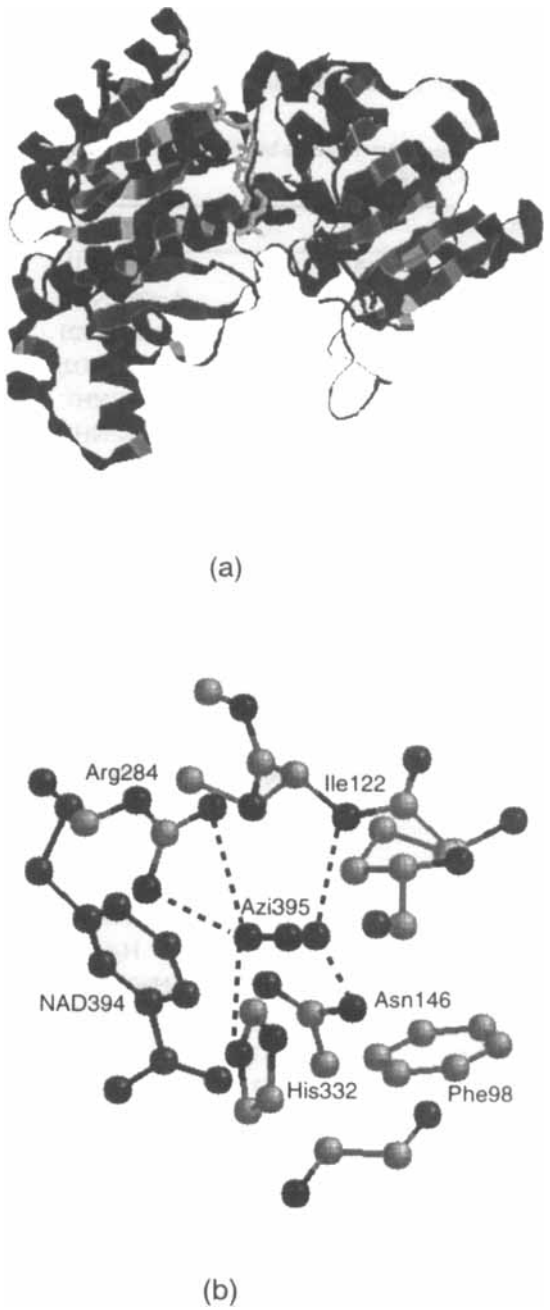


FIGURE 2 NAD-depend formate dehydrogenase complex (2NAD): a) Azide anion (blue colour) is located near the point of catalysis (catalytic site) in the proximity of NAD (yellow colour); b) Environment of azide. Hydrogen bonds are indicated with dashed lines (See Color Plate VI at the back of this issue)

There are two hydrogen bonds in each of independent molecules of **1MNG**: N3... HO (Tyr36) and N3...H_2O. In **1ISC** the terminal nitrogen atom of azide is H-bonded with His31 (N3... HN of 3.02 Å) and in **A** molecules there is a water molecule at the distance of 3.08 Å from the azide N1 atom coordinated with iron.

In myoglobin structures only terminal nitrogen atoms are involved in H-bonding: N3... HN (His64) of 2.91 Å in **1SWM** and N3...NH_2(Arg66) of 2.79 Å in **5MBA**.

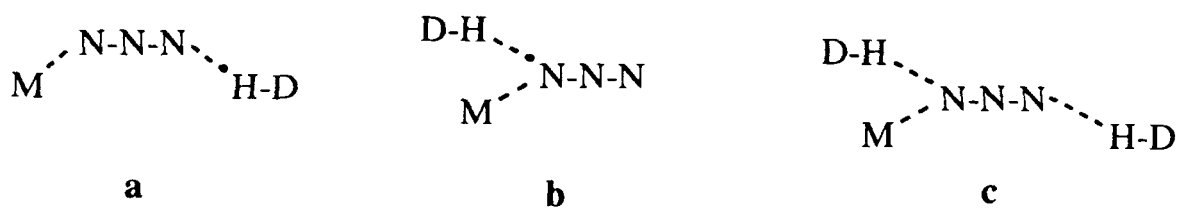
In the trinuclear copper cluster of ascorbate oxidase **1ASQ** some water molecules in the proximity of the azide ligand with the N...H_2O distances varying from 2.53 to 3.38 Å were reported.

In proteins **2HMZ, 1VNC, 2MHR** and **2TSB** the azide unit is in exclusively hydrophobic environement, and forms a short contacts (less then 3.4 Å) with the carbon atoms of the cyclic systems (Phe, His, Trp or Tyr).

This preliminary analysis of anion sites in proteins with the known crystal structures shows that in some cases the interpretation of observed results is not straightforward. Anion coordination which is based on the deposited coordinates, shows that the atoms used to satisfy the electron density calculated from the observed intensities do not give a chemically reasonable description of anion site geometry. This is best illustrated by the two above mentioned proposed models of azide coordination in trinuclear copper complex of ascorbate oxidaze. The large cavity limited by three copper atoms of the cluster seems to be sufficiently large to accommodate another type of azide binding. At least one azide can be bound to two copper atoms as a bridge. Analysis of hydrogen bonding is the most complicated part of the protein structure study. The irregularity of azide environment and considerable variations in in bond distances between the azide atoms and donor groups in the asymmetric units of proteins are observed (Table IV). Thus the limited number and sometimes rather low accuracy of the crystal structures for anion-containing proteins does not allow us to carry out a systematic and detailed analysis of real patterns in anion interaction geometry using only protein structural data. It turns out however, that such analysis may benefit substantially through the extensive use of small molecule structural information.

2. Metal-anion interaction in small molecule structures

This section deals with the results of the systematic analysis of the geometrical characteristics of the azido, thiocyanato and isothiocyanato functional groups and their immediate chemical environments performed with the help of the Cambridge Structural Database (CSD, [4]). The main purpose of the study is the detailed description of structural metrics relationships in free and covalently bonded anionic entities based on the atomic coordinates retrieved from the small molecule structures deposited in CSD. Some of the results (mostly concerning metal-thiocyanate interactions) were reported earlier [22].



SCHEME 4 Modes of H-bonding in M-NNN fragment

TABLE V Number of SCN fragments with various modes of metal complexation. Definition of interaction modes is given in Scheme 1

Metal/Mode	*A*	*D*	*E*	*F*	*G*	*N total*
		1a:				
Li	**21**	0	3	0	1	25
Na	**18**	0	1	0	1	20
K[a]	**21**	3	3	0	6	33
Rb[a]					0	3
Cs[a]					1	4
		2a:				
Mg	**5**	0	0	0	0	5
Ca	**14**	0	0	0	0	**14**
Sr	**7**	0	0	0	0	7
Ba	**6**	0	1	0	0	7
		3a:				
Al	2	0	0	0	0	2
Tl	1	2	0	0	0	3
		1b:				
Cu	**89**	**8**	0	3	36	118
Ag	**0**	**6**	0	3	8	13
Au	**0**	**7**	0	0	1	8
		2b:				
Zn	**21**	0	0	0	0	21
Cd	**12**	1	1	0	10	24
Hg	0	**17**	0	0	8	25
		6b:				
Cr	**54**	0	0	0	2	56
Mo	**153**	0	0	0	2	155
		7b:				
Mn	**36**	0	0	0	2	38
Tc	4	0	0	0	0	4
Re	**13**	0	0	0	0	13
		8:				
Fe	**49**	1	0	0	1	51
Ru	2	0	0	0	2	4
Os	0	0	0	1	0	1
Co	**94**	3	0	0	1	98
Rh	**0**	4	0	0	0	4
Ir	0	2	0	1	0	3
Ni	**188**	0	1	0	8	197
Pd	**16**	**17**	0	0	5	38
Pt	1	**8**	0	0	0	9

Notes: The digits in bold characters, represent the number of structures considered for the geometrical analysis.

a. Cations prefer to interact with π-system of thiocyanate, (**B** and **C** modes of coordination, Scheme 1).

TABLE VI Number of metal-azide fragments with various modes of complexation. Definition of interaction modes is given in Scheme 2

Metal/Mode	*A*	*B*	*C*
Cu	**73**	**47**	**20**
Mo	**36**	2	0
Co	**32**	0	1
Ni	**22**	**20**	**16**
Mn	**19**	1	1
Fe	**11**	0	0
Zn	3	**10**	0
Sum[a]	196	80	38
All[b]	270	109	50

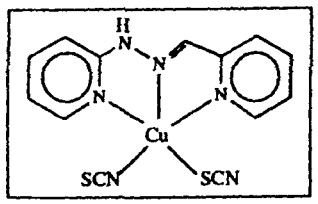
Notes: The digits in bold characters, represent the number of structures considered for the geometrical analysis.

a. Only fragments which occerences number is more 10 were considered;
b. Total number of retrieved fragments. All metals were considered.

2.1. *Modes of coordination*

We probe metal-anion interactions within several groups of elements of different size, charge and valence state. The following metals were analysed: alkali and alkaline-earth metals; the zinc and copper metal groups, transition metals; elements of group 7, trivalent metals. The search for small molecule crystal structures containing the metal-anion group resulted in 251 structures with azide and more than 850 structures with thiocyanate. Preliminary structural data analysis shows that all modes of anion coordination are realised; however the rates of their occurrences differ considerably.

The preferred types of interaction of the SCN with the metal cations are unidentate either through nitrogen or sulfur atoms (Table V, **A** and **D**), or bridging (most favoured **G** type) (Scheme 1) coordination modes. In 80% of all reported cases the metal is bound through the terminal N atom, in 9% through the S atom and in the remaining 11% the SCN entity acts as a bidentate ligand, i.e. it is coordinated at both ends. The azide also shows a preference for binding to a single metal center (62%) (Table VI).



SAHDOU

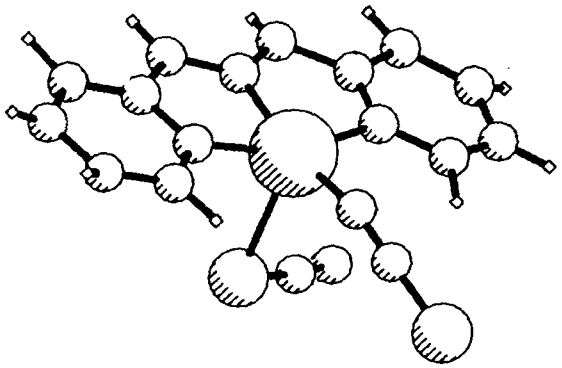
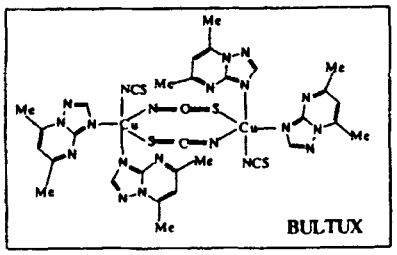


FIGURE 3 Pluto diagram illustrating coexistence of two different modes of coordination (Cu-SCN and Cu-NCS) in **SAHDAU**

However the rate of occurences of azides bonded to two metal centres is significantly higher (25%) than for SCN. Azide is coordinated at both ends (mode **C**, Scheme 2) as well as the thiocyanate, in 12%.

Although the anions are coordinated to a broad range of metals, the major part of the structures features coordination to Ni, Mo, Co, Cu, Fe and Mn ions. Other metals are either much more rare or occur in trace (or ultratrace) quantities.

Thiocyanate binds with Ni, Mo, Co, Cr, Fe and Mn through nitrogen. In a few cases (Cu, Ag, Cd, Hg) its shows non-negligible bridging mode of binding. S-Coordination was observed only for Ag, Au, Hg, Pd and Pt complexes. The coexistance of two different modes of coordination in the same crystal has also been reported in a few cases (Figures 3 and 4). The peculiarities of thiocyanate binding with the alkali and alkaline-earth metals deserve a special mention.



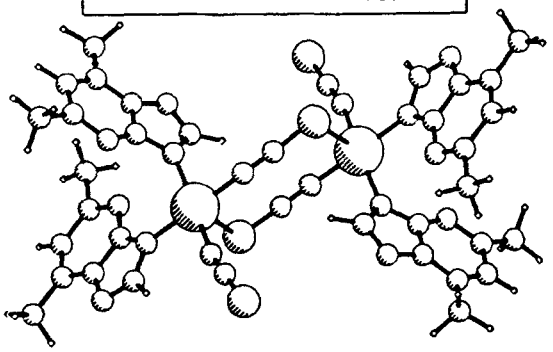
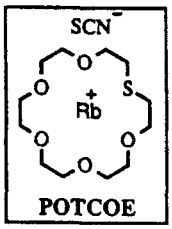


FIGURE 4 Pluto diagram illustrating coexistence of two different modes of coordination – by one terminal atom and/or forming the bridge-bonding in **BULTUX**

Whereas the lighter elements bind neatly in the **A** mode, heavier metals (M=K, Rb and Cs) prefer to interact with the π-system of thiocyanate (modes **B** and **C**) with no distinct acceptor S or N atoms (Figure 5).



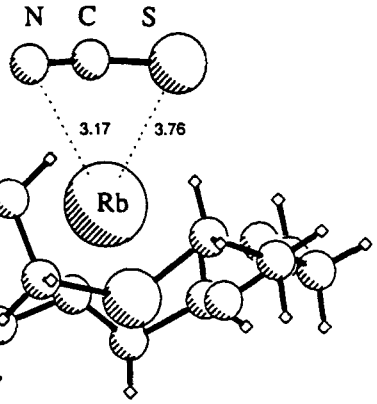


FIGURE 5 Pluto diagram illustrating lateral mode of coordination in **POTCOE**. Distances Rb...S 3.76, Rb...C 3.20, Rb...N 3.17 Å and bond angles ∠(RbSC) 58, ∠(RbCS) 96, ∠(RbNC) 81° show **B** or **C** mode

TABLE VII Mean values of the geometrical parameters with mean SD for the geometry of differently substituted azides. Definition of interaction modes is given in Scheme 2

Mode/Parameter	*NNN*	*A*	*B*	*C*	*R*[a]
N1-N2	1.158(5)	1.175(2)	1.202(5)	1.170(4)	1.220(3)
N2-N3	1.159(5)	1.151(2)	1.149(7)	1.171(3)	1.130(2)
N1...N3	2.317(7)	2.325(3)	2.351(7)	2.370(5)	2.344(3)
∠(N1N2N3)	179.0(3)	176.0(2)	178.0(3)	177.6(2)	172.5(2)
M-N1		2.050(8)	2.12(2)	2.18(2)	1.51(1)
M-N3				2.18(3)	
∠(MN1N2)		126.4(5)	124.7(7)	131(2)	115.9(2)
∠(MN3N2)				131(2)	
N obs	19	270	109	50	208

a. R = C, B, P, Si and I in R-NNN compounds

TABLE VIII Mean values of the geometrical parameters with mean SD for M...NNN. (Nitrogen cov. rad. = 0.70 Å)[a]

Metal	*Nobs*	*M...N*	*<(MNN)*	*sum of cov.*
Cu	73	1.996(9)	126.1(9)	2.05
Mo	36	2.09(1)	128(1)	2.10
Co	32	1.964(6)	122.3(8)	2.02
Ni	22	2.06(2)	128(1)	2.09
Mn	19	2.07(3)	129(3)	2.07
Fe	11	2.001(2)	127(1)	1.93
Zn[b]	10	1.99(1)	125(1)	2.01

a. All values for covalent and ionic radii from Wells [23].
b. Zn-NNN-Zn complexes were considered.

Azide shows a preference of binding in **A** mode with Mo, Co, Mn and Fe cations. There is a definite preference for binding twice through one terminal nitrogen (**B** mode) in case of the Zn complexes. With two cations, Cu and Ni, azide shows all modes of coordinations.

Metal complexes with thiocyanate or azide exhibit several coordination numbers and a variety of metal oxidation states. Cu and Ni have a oxidation state of II in all cases with coordination numbers ranging from 4 to 6. Co has oxidation states II and III with coordination numbers 4, 5 and 6. Fe may have oxidation states II and III and coordination numbers 5, 6 and 7. Mn has oxidation states II, III and IV and coordination numbers in the range from 4 to 7.

2.2. *Geometrical characteristics of metal-anion binding*

In order to characterize the metal -anion interactions we selected only the unidentate fragments of type **A** and **D** for thiocyanate (Schemes 1) as all other cases cover just a small portion of all structures. In azide complexes all modes were considered.

Each mode of azide binding is associated with the characteristic values of M-N1 distance and ∠(MN1N2) angle. Thus for the **A**, **B** and **C** modes these values are 2.050(8), 2.12(2), 2.18(2) Å and 126.(4)°, 124.7(7) and 131(2)°respectively (Table VII). Further subdivision of structures according to the type of cations indicates that the bond parameters vary very slightly (Table VIII). We did not find any obvious correlation betwen the metal-azide geometry metrics and metal oxidation state or coordination number.

The different situation is observed in thiocyanate complexes. The mean values of bond distance M...N and valence angle ∠(MNC) for each metal are listed in Table IX. The M-NCS fragment geometry shows larger variations in the M...N distance as well as significant variations in the angles at the N atom.

TABLE IX Mean values of the geometrical parameters with mean SD for M...NCS. (Nitrogen: ionic rad. = 1.46 Å, cov. rad. = 0.70 Å)[a]

Metal	*Nobs*	*M...N*	*<(MNC)*	*sum of ionic radii*[a]	*sum of cov. radii*[a]
1a:					
Li	22	1.987(7)	168(2)	2.06	
Na	20	2.38(1)	150(3)	2.41	
K	28	2.85(2)	133(4)	2.79	
2a:					
Mg	10	2.093(8)	169(1)	2.11	
Ca	31	2.437(8)	160(2)	2.45	
Sr	18	2.624(5)	161(2)	2.59	
Ba	13	2.84(1)	152(4)	2.81	
1b:					
Cu(all)	135	2.03(1)	162(1)		2.05
Cu(clust)[b]	104	1.944(5)	166.1(6)		
2b:					
Zn	33	1.99(1)	163(4)		2.01
Cd	24	2.28(2)	152(3)		2.18
6b:					
Cr	54	1.983(3)	170.6(8)		1.99
Mo	153	2.124(5)	168.5(5)		2.14
7b:					
Mn	55	2.04(1)	171(1)		2.07
Re	41	2.163(8)	160(1)		
VIII					
Fe	74	2.043(7)	166(1)		1.93
Co(all)	185	1.983(5)	167.8(6)		2.02
Co(III)	120	1.934(3)	169.4(6)		
Co(II)	65	2.074(4)	165(1)		
Ni(all)	301	2.043(4)	165.8(5)		2.09
Ni(4,5)	38	1.893(9)	171.1(9)		
Ni(6)	263	2.065(2)	165.1(5)		
Pd	15	2.027(7)	166(1)		2.11

a. All values for covalent and ionic radii from Wells [23].
b. The cluster corresponds to the most populated area.

The binding of thiocyanate anion with the alkali and alkaline-earth metals (**1a** and **2a**) is essentially ionic. The size of the cation may influence its disposition relative to the delocalized thiocyanate (Figure 6): the smaller cations (Li^+, Na^+ and Mg^{+2}) clearly interact solely with nitrogen, whereas the larger ones (e.g., K^+ and Ba^{+2}) prefer to interact with a larger part of the linear anion (π-system of SCN). Thus, it appears that there is a correlation between the ionic radius and the M-NCS angle (Figure 7).

The interactions with other metals show a predominantly covalent character. Comparison of the metal-thiocyanate geometry for different metals within the same group shows that the larger cations prefer interacting with many centres (π-system of C-N bond) as was observed for alkali and alkaline-earth-metals. Thus, the cations Zn and Cd or Mn and Re show different positions above the delocalized SCN (Figure 8). The geometrical parameters of Zn...NCS [mean 1.99(1) Å and 163(4)°] differ from those for Cd [2.28(2) Å and 152(3)°]. The mean values of bond and angle distances Mn...NCS and Re...NCS are 2.04(1) Å, 171(1)° and 2.163(8) Å, 160(1)°.

Among the group **8** metals the thiocyanate complexes were reported only for Fe, Co, Ni, and Pd; all of them are M-N-bonded. Two clusters in the cobalt complex correlation diagram (Figure 9, a) correspond to different oxidation states of Co: the larger Co-N distance [2.074(4) Å] indicates a six-coordinated Co(II) complex while the shorter Co-N distance [m.v. 1.934(3) Å] is usually observed for 4- or 5- coordinated Co(II) and for 6- coordinated Co(III). The geometry of Ni-complexes shows a good example of geometry *versus* coordination number correlation (Figure 9, b). The small cluster (38 observations) [1.893(9) Å, 171.1(9)°] corresponds to 4 or 5 coordinated Ni(II), whereas the second cluster (263 observations) with the larger distance [2.065(2) Å] and valency angle [165.1(5)°] corresponds to 6-coordinated Ni(II) only. This is in accordance with the well-known fact, that the covalent radius for a given cation increases together with the coordination number [23].

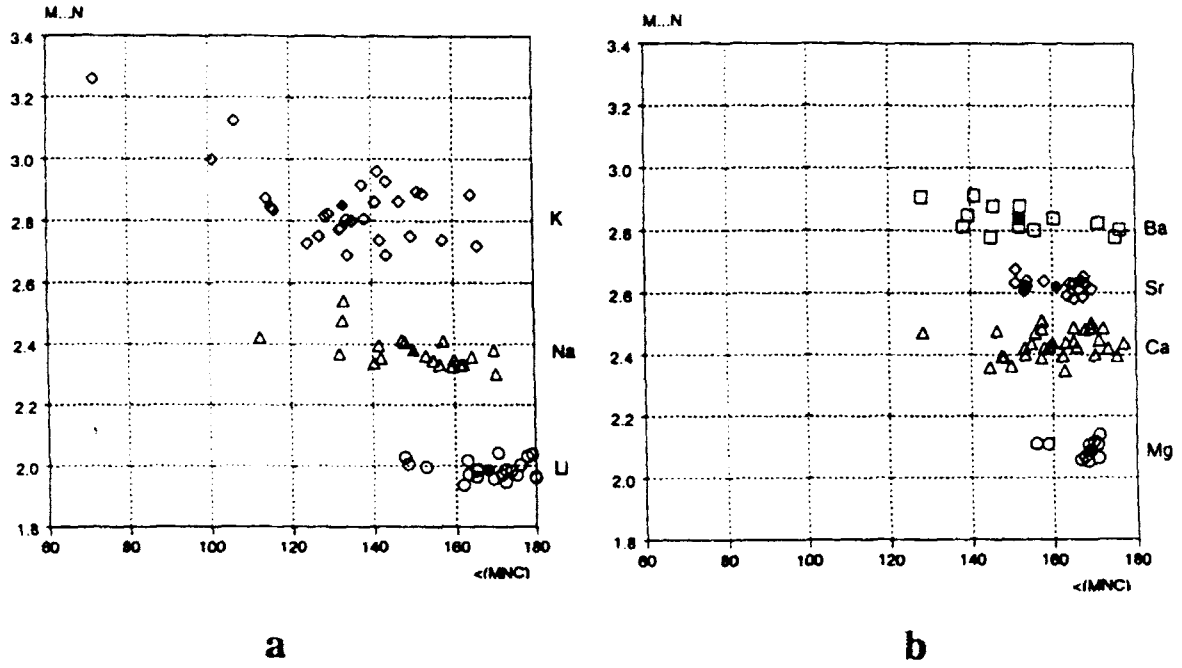


FIGURE 6 Scatterplots of distances M...N *versus* ∠(MNC). (a) – alkali metals; (b) – alkaline earth metals. The circles are Li (a) and Mg (b); triangles are Na (a) and (Ca); rhombs are K (a) and Sr (b); squares are (Ba). Black symbols correspond to the mean values for each metal

TABLE X Average value of M...SCN geometry (with mean SD) (sulfur: ionic rad. = 1.84 Å, cov.rad. = 1.04 Å[a])

M	*Nobs*	*M...S*	*<(MSC)*	*ionic radii*	*cov. radii*
Cu[a]	16	2.57(7)	98(2)	2.80	2.39
Ag	10	2.53(8)	102(2)	3.10	2.57
Au	12	2.54(8)	102(2)	3.21	2.54
Hg	77	2.507(9)	99.0(4)	2.94	2.52
Pd	25	2.351(7)	106.2(8)		2.41
Pt	13	2.36(2)	106.4(9)		2.43

a. All values for covalent and ionic radii from Wells [23].

The geometry of M...SCN fragments shows that M...SCN distances are in the range from 2.351(7) Å (Pd) to 2.57(7) Å (Cu) (which corresponds to the predominantly covalent bonding), and the valence angle ∠(MSC) is characterized by a relatively small dispersion [98(2) – 106.4(9)°] (Table X). We can distinguish only two different clusters belonging to Hg and Pd, Pt (Figure 10). The discrete tail above 2.6 Å corresponds to interactions with the partial ionic character for some metals such as Ag, Au, Cu [23].

2.3 *Structural metrics relationships of covalently bonded azide and thiocyanate*

The geometry and electronic structure of SCN as well as its thiocyanato and isotiocyanato derivatives have been studied in a series of experimental [24] and theoretical papers [25]. Theoretical aspects and the electronic structure of covalent azides have been reviewed earlier by Treinin and Kaftory [26] and more recently by Klapotke [27]. The structure of azido- and thiocyanato-groups in solid state was also briefly covered in a review of triply-bonded functional groups [28]. The large number of observations we collected allows us to analyze the influence of covalent bonding on the internal geometry of azide and thiocyanate. We retrieved from CSD all struc-

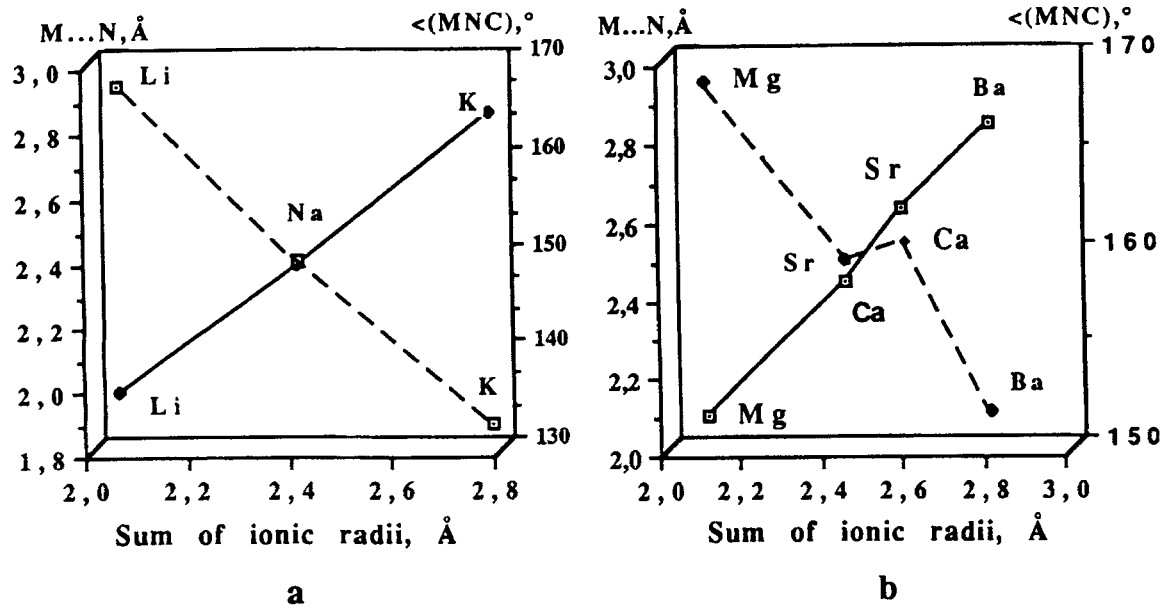


FIGURE 7 Correlation of the bond distances (solid line) and angles (dashed lines) *versus* ionic radii: (a) – alkali metals and (b) – alkaline earth metals

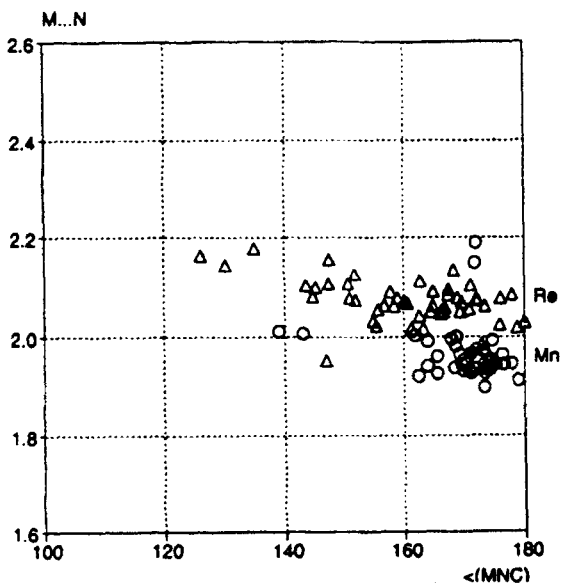


FIGURE 8 Scatterplots of distances M...N *versus* ∠(MNC) in the metal complexes of the **7b** group. The circles are Mn and triangles are Re. Black symbols correspond to the mean values for each metal.

tures with the NNN and SCN fragments bonded to a metal or a non-metal atom as well as all these units in the anionic form.

The average dimensions of the non-bonded SCN are: S-C 1.624(4), C-N 1.137(5) Å, ∠(SCN) 175.7(4) (Table XI). Taking into account the experimental accuracy, we assumed SCN group to be perfectly linear thus complying with the $C_{\infty v}$point-group symmetry. In covalently bonded thiocyanate we find practically no variation in the C-N bond length and SCN bond angle regardless of the binding mode. But the S-C bond is much more sensitive to both the mode of coordination (by sulfur or by nitrogen) and the nature of the partner. The extreme values [the shortest in SCN-R 1.574(5) Å and the longest in R-SCN 1.683(5) Å] belong to the bonds with a non-metallic ligand ("pure covalent"). This peculiarity may be explained by different contributions of the limiting canonic forms in both cases. Thus, the bonding in the SCN-R system is determined primarily by the charge-smeared

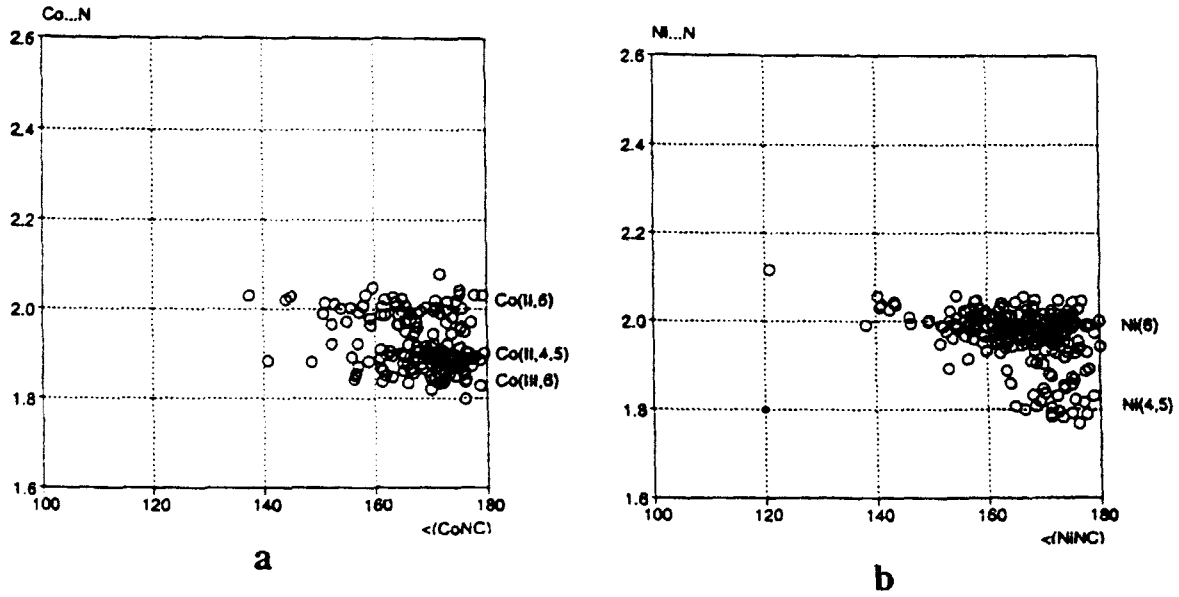


FIGURE 9 Scatterplots of distances M...N *versus* ∠(MNC) in the metal complexes of the **8** group. Black symbols correspond to the mean values for each metal

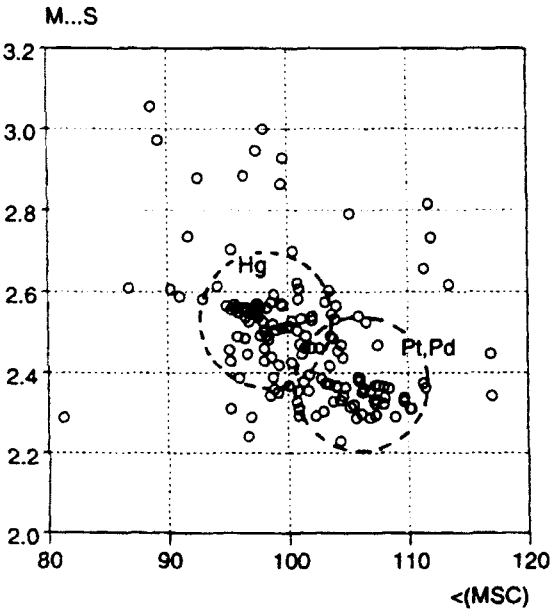


FIGURE 10 Scatterplots of distances M...S *versus* ∠(MSC). The preference areas corresponding to Hg and Pt, Pd are shown

S=C=N-R and $^{+}$S≡C-N^{-}-R components rather than by the $^{-}$S-C≡N^{+}-R form, whereas the R-SCN mode of coordination, on the contrary, features a greater contribution of uncharged R-S-C≡N bonding pattern than that of the R-S^{+}=C=N^{-} mesomer. Similar behaviour of the S-C bond length is observed in metal complexes, although the difference in these cases is less obvious [1.621(1) Å in SCN-M and 1.648(4) Å in M-SCN].

The azide in the ionic form has a $D_{\infty h}$ symmetry, with a mean value of the ∠NNN 179(2)° and N-N bond of 1.158(4) Å. This corresponds well to the proposed resonance between the three canonical structures (**Ia**) – (**Ic**) [23]: (**Ia**) $^{-}$N=N^{+}=N^{-}, (**Ib**)$^{--}$N-N^{+}≡N and (**Ic**) N≡N^{+}-N^{--}. As well as thiocyanate, the azide unit is nearly linear in all modes of binding with the metal (Table VII). The average values of ∠(N1N2N3) are 176.0(2) in **A** mode, 178.0(3) in **B** and 177.6(2)° in **C**. In both unidentate modes **A** and **B** the difference between the N-N distances is not negligible, 0.024 Å for **A** mode and 0.05 Å for **B** mode. Com-

parison of these distances with the metric of discrete azide shows clearly that the N1-N2 distances are elongated by 0.02 –0.05 Å and the N2-N3 distances are shortened by 0.01 Å. In the bridging mode **C** both N-N bond distances are identical, but both of them are elongated by 0.02 Å as compared to the typical value for non-bonded azide. In the azide covalently bonded to non-metallic atom the distortion of N-N-N unit is more evident, with the ∠ N1-N2-N3 of 172.6(2)° and with two significantly different ($\Delta = 0.1$) N-N bond lengths, N2-N3 of 1.128(2) Å and N1-N2 of 1.222(3) Å. The relationships in these triatomic fragments exhibit that the triple-bond character of the terminal N2-N3 bond increases as the single-bond character of N1-N2 decreases which may be associated with the non-uniform contributions of the limiting canonic forms, that of (**2b**) being obviously overwhelmingly higher than that of (**2c**): (**2a**) R-N=N^+=N^-, (**2b**) R-$^-$N-N^+≡N, (**2c**) R-N^+≡N^+-N^{--} [23].

The covalent binding significantly influences the values of all bond parameters. The equivalent N-N bond distances and linearity of non-bonded azide indicate that the most important factor in the system is the π-delocalisation over the triatomic unit (*i.e.* resonance). On the other hand, in covalently bonded azide there is a competition between resonance and a strong negative (anionic) hyperconjugation [29] which donates electron density from the filled σ (R-N1) orbital into the unfilled, antibonded π^*(N2-N3) orbital. Negative hyperconjugation is important in charged and neutral systems as well. Substitution effect reduces the azide unit symmetry from a $D_{\infty h}$ to C_S.

3. How metal cations can modulate the hydrogen bonding of thiocyanate and azide

The role of the metal cation in the three-dimensional H-bond network is unclear. It may act as a template, dictating the orientation of the remote hydrogen bonding surfaces as a consequence of the stereochemical preferences of the cation. On the other hand, the metal cation may exert a specific electronic influence on the groups at the remote hydrogen bonding site, thereby providing a means for tuning the strength of intermolecular interactions. To be able to discuss the H-bonding capability of the metal bonded anions, we have to compare it with that in the discrete anions. Our next immediate task is to present a survey of the molecular recognition properties of these functional groups derived from their non-bonded interactions in the solid state.

TABLE XI Comparative analysis of the internal geometry of differently substituted thiocyanates

fragment/parameter		*S-C*	*C-N*	*<(SCN)*	*Nobs*
SCN-R[a]		1.574(5)	1.153(5)	176.8(3)	42
SCN-M	(all)	1.621(1)	1.146(1)	177.0(9)	1724
	(cov)[b]	1.621(1)	1.150(1)	177.5(1)	905
	(ionic)[c]	1.626(3)	1.147(2)	177.(1)	155
Discrete SCN		1.624(4)	1.137(5)	175.7(4)	258
M-SCN		1.648(4)	1.152(3)	176.4(4)	227
R[a]-SCN		1.683(5)	1.147(4)	176.0(4)	30

a. R = C, B, P, Si and As in SCN-R and C in R-SCN compounds;
b. - only covalent M...NCS interactions
c. - only ionic M...NCS interactions.

TABLE XII Multiple H-bonding of thiocyanate and azide in small molecule structures

H-bond/ Acceptor	*SCN*		*M-SCN*	*M-NCS*	*NNN*	*M-NNN*
	S	*N*	*N*	*S*	*N*	*N*
1 HB	34	36	9	103	7	18
2 HB	20	39	6	47	2	9
3 HB	11	15	3	24	5	1
4 HD	21	4		10	3	
> 4HD	2			3		
total	88	93	18	187	17	28

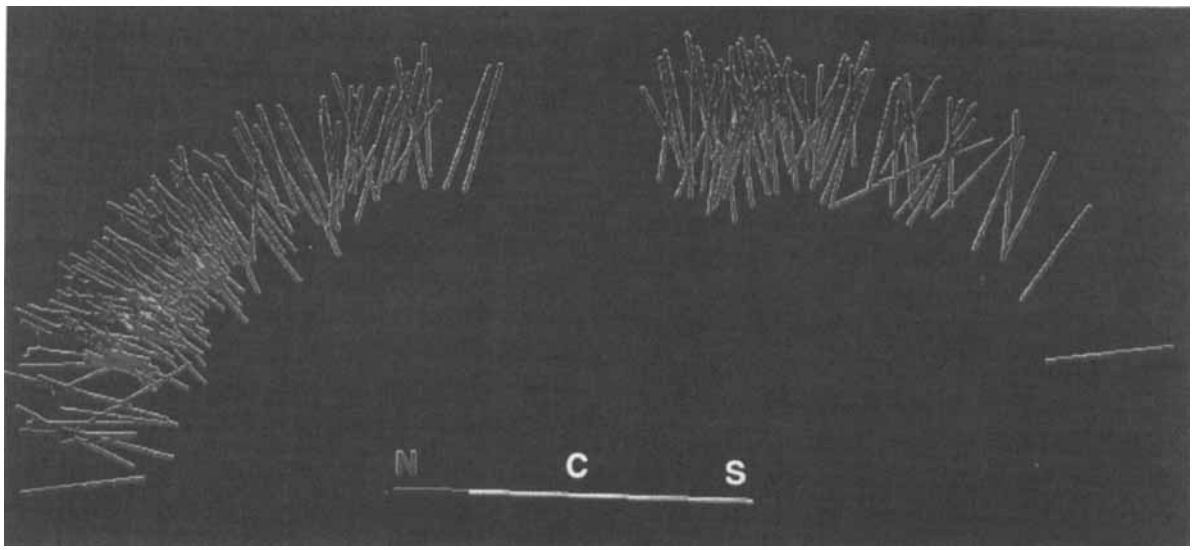


FIGURE 11 Graphical representation of the preffered areas for H-donor groups distribution around the S- and N-acceptors. N-H and O-H donors were merged together (See Color Plate VII at the back of this issue)

The structures with discrete azide which are also rich in donor groups are quite rare. However the number of relevant structures with discrete thiocyanate is sufficient for reasonable analysis. Both anions appear to show a similar accepting behaviour: in fact both are able to accept vitually all types of hydrogen donors, O-H and N-H groups forming strong hydrogen bonding. As well as metal cations, the donor groups (D) can interacts with the anions either through the terminal atom (acceptor center A) or interact with the π-system of the anion. Each of the anions is also capable of binding *via* both terminal atoms, thus forming a bridge. It was found that in most cases (80%) anions bind to donor group through both acceptor centers. An acceptor center of the anion forms a total of one, two or three hydrogen bonds.

Each acceptor centre of SCN participates in a single hydrogen bond only in 38% of all cases; in 62% of all structures at least one of the acceptor centres is engaged in multiple H-bonding: in most cases, the nitrogen atom binds twice (39 occurrences) but three and four H-bonds are also observed in some structures (15 and 4 cases, respectively; see Table XII) [30]. The occurrences of two, three and four H-bonds around the sulfur are observed in 20, 11 and 21 cases, respectively. In three fragments sulfur shows five- or even six-fold coordination with H-donor groups. Obviously, the discrete anion forms an almost equivalent number of H-bonds at both ends. The relative orientations of the donor groups are different around sulfur and nitrogen (Figure 11).

The observed preference of N-metal coordination leads mainly to data for H-bonding through sulfur. For metal-bonded isothiocyanate we encountered 309 H-bonds from 184 terminal sulfurs and only 30 H-bonds from 18 terminal nitrogens for thiocyanate. Very small number of SCN groups bonded to the metal atom through sulfur is not sufficient for a detailed statistical study of the accepting properties of terminal nitrogen. Consequently, we shall limit the discussion to the comparable statistics on sulfur as H-acceptor. However, the details concerning the H-bonding

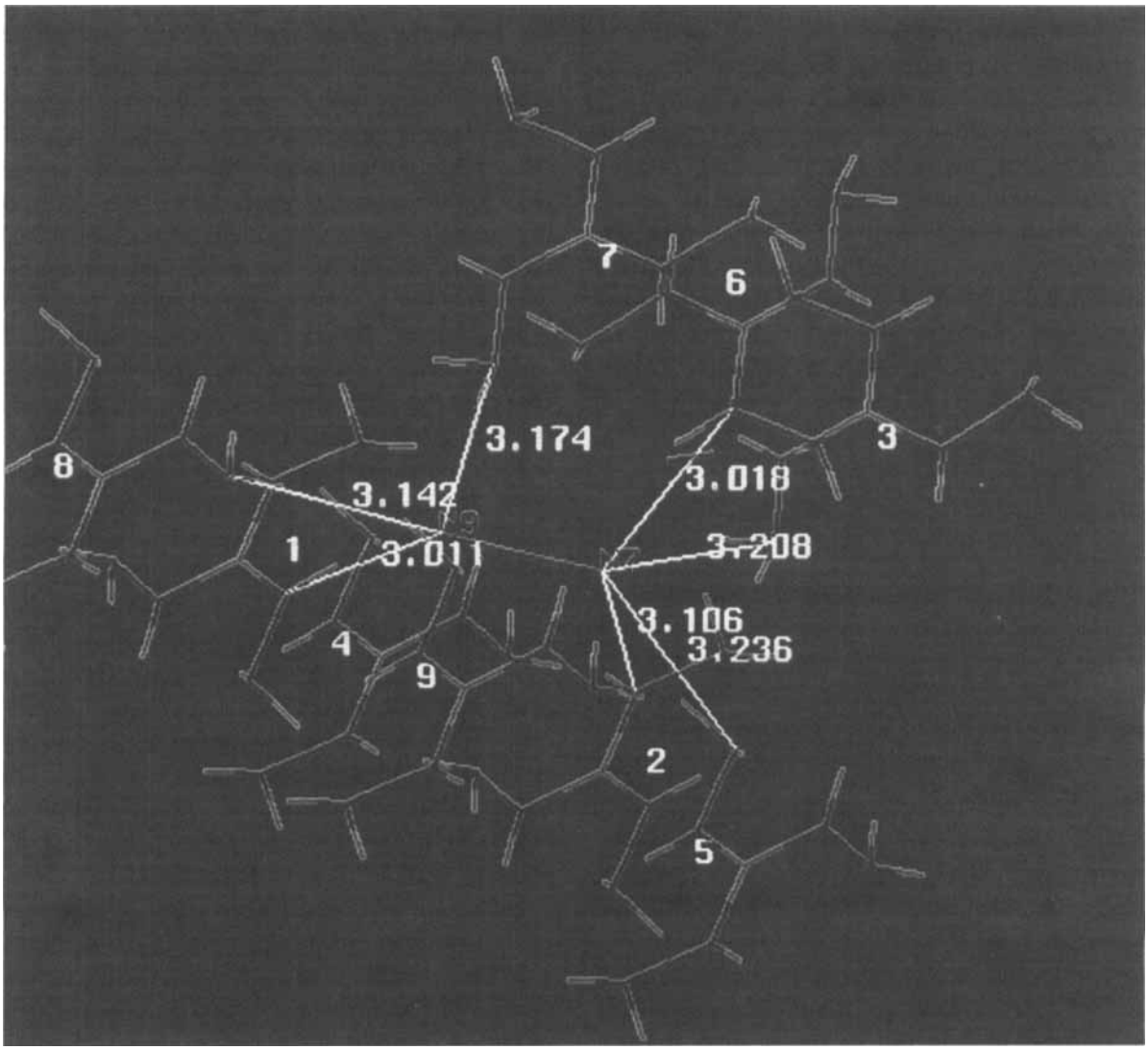


FIGURE 12 H-bonding of the azide in **KESVOT**. DH...N distances (Å) are shown. Numbering indicates the symmetry-related molecules (See Color Plate VIII at the back of this issue)

of nitrogen will be illustrated using selected structures.

Table XII shows that the multiple H-bonding involving sulfur occurs much more rarely for M-NCS than for the discrete anion. A single hydrogen bonding is observed in 55% of all occurrences when SCN is metal-bonded (compared to 38% for the discrete anion). The strength of the hydrogen bond is only slightly influenced by the metal linked to the anion [30].

Discrete azide also participates in H-bonding through both terminal nitrogens, in the structures rich in H-donor groups it normally forms multiple hydrogen bonds (Table XII). Number of H-bonds vary from 2 to 7. A good example of an extensive H-bonding pattern around the azide anion in the complex with 1, 2, 3-triaminoguanidinium (**KEZVOT**) is shown in Figure 12.

In azides bonded to metal the metal-coordinated nitrogen usually forms no more than a sin-

gle H-bond; the terminal N atom is usually involved in one or two H-bonds. Considered structures of covalently bonded azide with non-metal atom are rich with donor groups (70 structures contain one or more donor groups but in most cases the potential H-bonding partner prefers another acceptor group such as carbonyl). It was found, that the terminal nitrogen atom of a covalently bonded azide participates in hydrogen bonding only in a few structures (**DAHLIH**, **FIZDIB** and **VIRZUR**) forming a very weak H-bond with a hydroxy group: (N...O 3.10, N...H 2.74 Å and ∠ at H 109°) [31].

DISCUSSION

From a chemical point of view, all types of interactions realized in small molecules are identical to those observed in macromolecules. Thus it appears natural to compare the observations obtained by statistical analysis of structural fragments in small molecule structures to those in macromolecules.

Both anions exhibit a rich variety of binding modes with their partners, metals or H-donor groups. In discrete form both of them show an extremely high H-bond acceptor properties with respect to OH or NH groups. Free anions, SCN and NNN, have intermediate dimensions which correspond well to the geometry with the net charge on both terminal atoms being the same, *i.e.* both terminal atoms are electron-rich and thus have excessive negative charge whereas the middle atom is electron-deficient and, hence, positively charged. This agrees with the observed identical ability of both terminal atoms to accept a donor group. Two proteins, **lysozyme** [7] and **erabutoxin-b** [8] containing the SCN in anionic form and the unique example of discrete azide in the structure of **NAD** [16] and also, a multiple small molecule structures, clearly demonstrate a non-covalent bridging function of these anions.

Both anions form strong complexes with metals. The preferred mode of binding is unidentate coordination. The bridging type of bonding (*i.e.* coordination at both terminal atoms) is realised in the same proportion for both anions (12% of all considered fragments). In unidentate mode SCN shows a preference to be bonded only to one metal, whereas azide in many structures binds twice (with two metal centres). The ambidentate function of thiocyanate gives rise to a linkage isomerism, *i.e.* to N-thiocyanato and S-thiocyanato complexes. The observations indicate that the SCN anion most frequently binds through the nitrogen atom.

Both anions are highly selective ligands. In the terms of "hard/soft" concept [32] we can conclude that azide binds preferably with hard (Mn^{2+}, Co^{3+}) and borderline (Zn^{2+}, Cu^{2+}, Ni^{2+}, Fe^{2+}, Fe^{3+}, Co^{2+}) cations. The selectivity of SCN is determined also by the different nature of terminal atoms. Thus it binds through the sulfur atom only with Hg, Pt, Ag, Au, Pd and Cu cations. The N-thiocyanato complexes are formed with alkali and alkaline-earth cations (those are mostly ionic interactions) as well as with the Ni^{2+}, Co^{2+}, Co^{3+}, Fe^{2+}, Fe^{3+}, Mn^{2+}, Mn^{3+}, Zn^{2+} and Cu^{2+} cations (covalent binding).

In the sea hare **myoglobin** (*Aplysia Limacina*) (ferric) complex with thiocyanate [11], the SCN anion was reported to coordinate the *haem* iron through the sulfur atom. As our analysis of small molecule complexes has shown, this looks like the case of realization of a low probability event (1 case out of 50).

Our analysis allowed to derive some quantitative characteristics of the metal-anion interactions. Thus in M-SCN fragments only variations of metal bond distances corresponding to the differences in the metal covalent radii appear to be meaningful, whereas the angular parameter is almost equal for all thiocyanate complexes (≈100°). The M-NCS fragment geometry shows larger variations in the M...N distance as well as significant variations in the angle at the N atom. This analysis reveals the correlation between oxi-

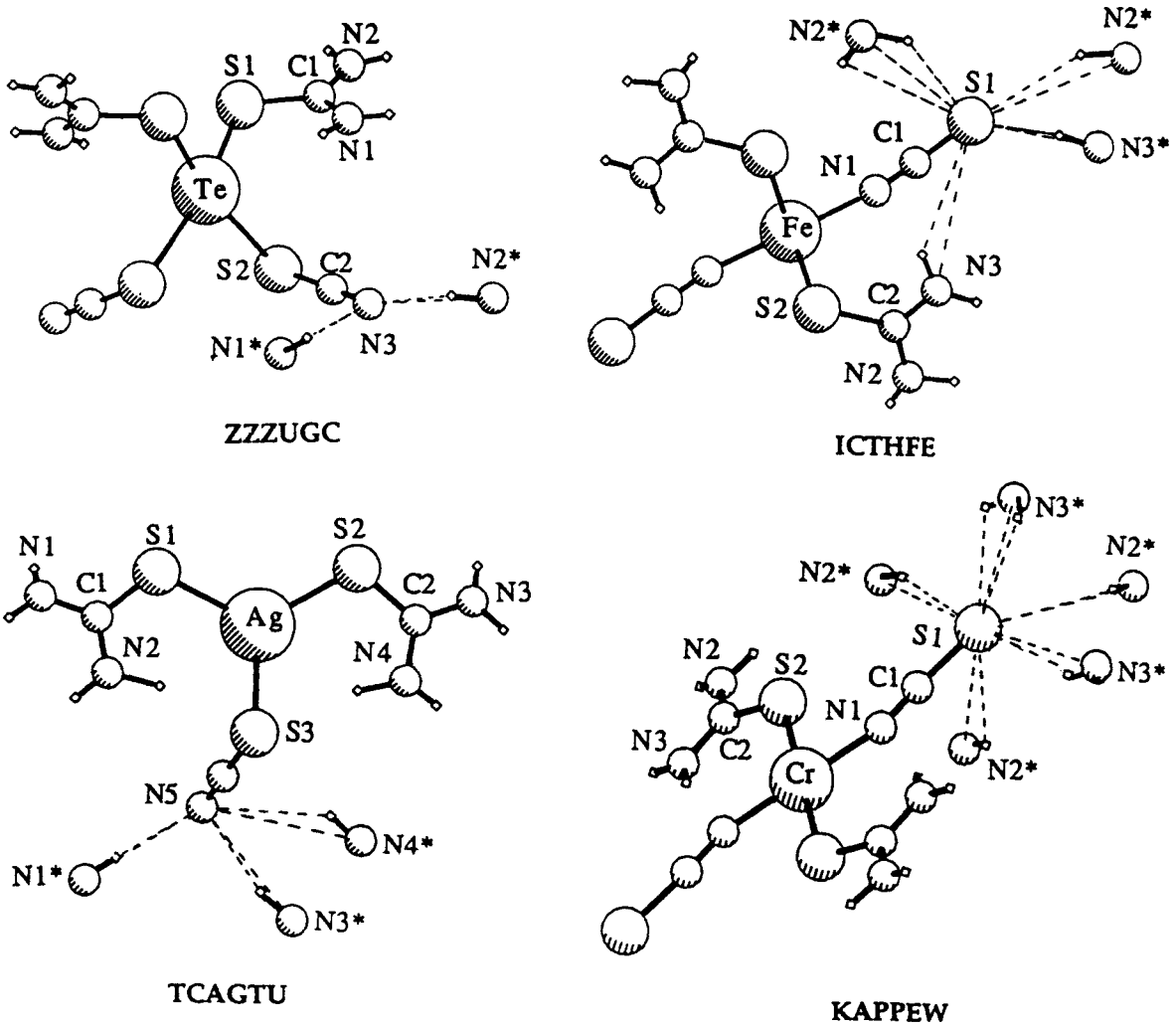


FIGURE 13 H-bonding of the terminal thiocyanate atoms in thiourea-family complexes: **ZZZUGC, TCAGTU, ICTHFE** and **KAPPEW**. H-bonds are shown by dashed lines for only crystallographically independent fragments

dation state and the covalent radius in the Co complexes and the correlation between coordination number and the covalent radius in the Ni complexes (Figure 9). In azide complexes the M...N bond distance and the MNN bond angle vary slightly for different cations.

The analysis of non-covalent interactions shows that the involvement of covalently bonded anionic group in multiple H-bonding is much less probable than in the case of discrete anions. This effect is certainly the result of metal coordination, which provokes the redistribution of negative charge and ultimately reduces the H-bond acceptor function of terminal atoms; it is observed both in small molecule and macromolecular structures. Excellent examples presenting the different modes of metal-complexation of thiocyanate and H-bonding with the same thiourea ligand are shown in Figure 13. M-SCN fragment forms only two H-bonds in **ZZZUGC01**

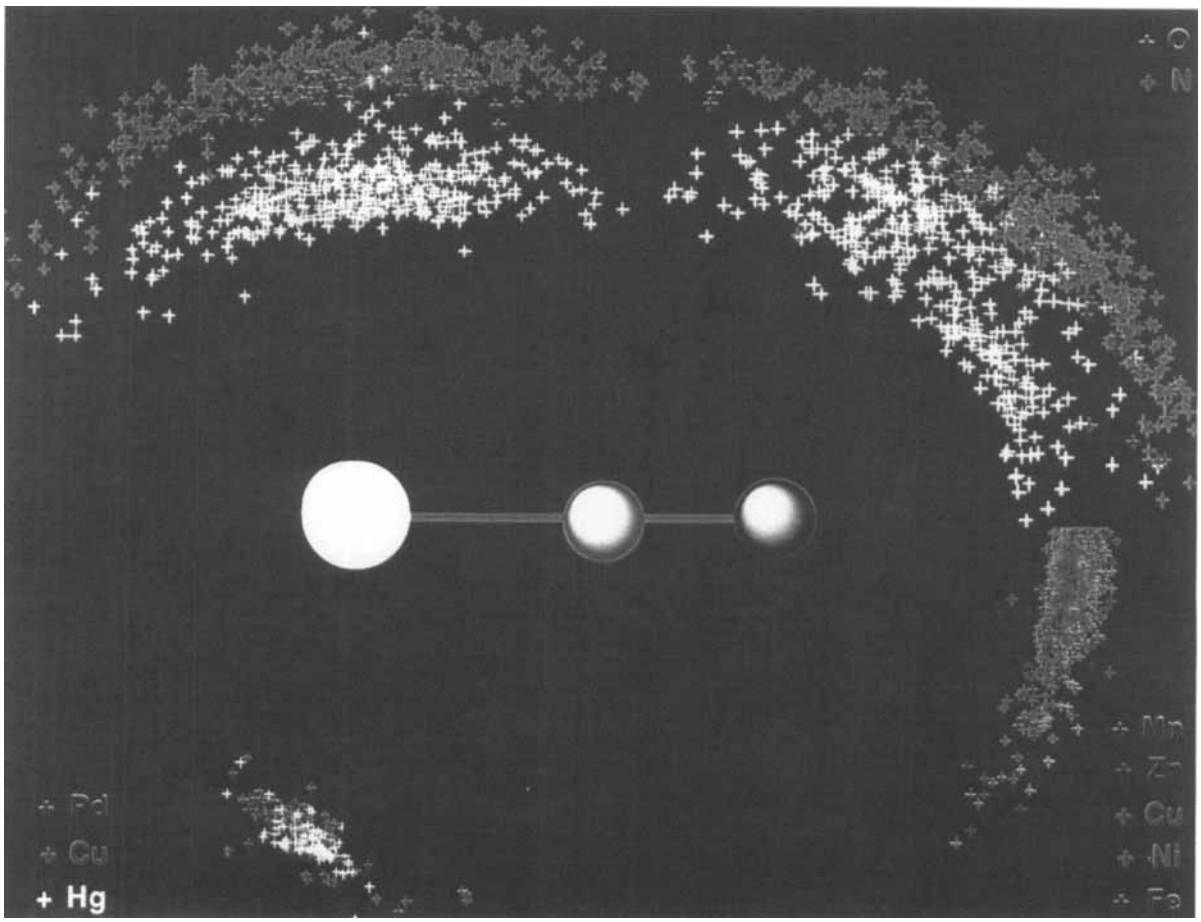


FIGURE 14 Graphical representation of the distribution of different type of partners (metals and H-donor groups) around the SCN. The areas of metal cations which interact with SCN through sulfur (on the left) and through nitrogen (on the right) are shown on the bottom of linear SCN. Hydrogen donor groups (on the top) are distributed in diffused domains around sulfur and nitrogen. The white crosses represet the hydrogen atoms, the blue and red crosses correspond to nitrogen and oxygen atoms (See Color Plate IX at the back of this issue)

and three H-bonds in **TCAGTU** through the terminal nitrogen atom while M-NCS fragment accepts, through sulfur, multiple hydrogen atoms from the five donor sites in **KAPPEW** and from the four donor sites in **ICTHFE**.

Analysis of anions environment in protein structures shows that the azide often binds to side chain structures (Arg, His, Asn, Thr and Tyr) and much more rarely forms H-bonds with the NH groups of the main chain. In some of the structures the anion environment is completed by the interactions with water molecules. The *azide-donor group geometry* are sifted to bring out any preferred direction in which azide holds a donor groups. Only in two cases the donor group is oriented along azide axis direction, the remaining 33 donor groups are considerably displaced off the azide axis, so that the mean value of ∠(DN3N2) angle is close to 114°.

All statistical results for both types of anion interaction (metal complexation and hydrogen bonding) are summarized in a graphical presentation of superimposed fragments (where all fragments are superimposed to form a giant "statistical molecule") (Figure 14 and 15).

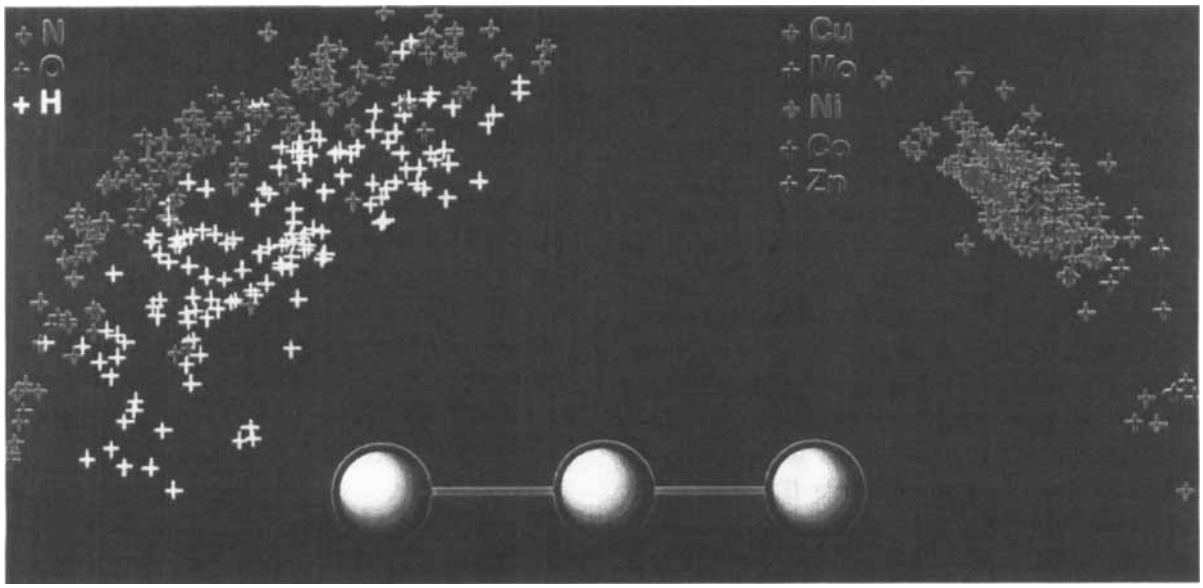


FIGURE 15 Graphical representation of the distribution of different type of partners (metal cations and H-donor groups) around the azide. The white crosses represent the hydrogen atoms, the blue and red crosses correspond to nitrogen and oxygen atoms (See Color Plate X at the back of this issue)

The points corresponding to coordinted metal positions in metal complexes are clustered along the preferred direction; the H-bond donor-groups localisations, conversely, form a diffuse arc around the anions. The metal cations bonded to the sulfur atom are clustered in a narrow angular domain with ∠(MSC) from 98 to 106°, in contrast to a much more extensive cluster at nitrogen which corresponds to the different metal positions with respect to the linear anion: the largest cations interact with the π-system of the anion (the bottom part of Figure 14). The dense metal cluster at azide nitrogen is localized in the range of M...N distances between 1.86 and 2.3 Å and M...N1N2 angles between 115 and 135° (Figure 15).

The relative orientations of the donors are different around sulfur and nitrogen for the thiocyanate (the top part of Figure 14). Although the distribution areas of donors around sulfur and nitrogen overlap, we can distinguish the principal binding zone of every acceptor centre: for sulfur-acceptor, there is a cluster of donors in a nearly perpendicular direction to the thiocyanate bar, whereas for nitrogen-acceptor, all donors gather mostly in the angular domain from 120 to 160°. The observed distribution of donor groups around azide is much more diffuse (Figure 15).

The role of cation is more important in the anion complexes but in covalently non-metal bonded anion the effect of atom charge localisation is much more evident; in particular, the delocalization has a dramatic influence on the H-bond acceptor properties of the anion. This may explain why we did not find the effective hydrogen bonding for both covalently non-metal bonded anions.

Quite recently, the first X-ray study of the protein complex of the azido-thymidine disphosphate (AZT) with the nucleozide diphosphate (NDP) kinase was investigated in crystal by Janin [33], in an attempt to understand why NDP kinase is so inefficient in AZT activating in comparison with the natural substrates. It was concluded that the dominant factor in the low activity of NDP kinase on the substrate (AZT) is associated with the lack of azido group H-bonding. Evidently, the above results are in a good accordance with these observations.

METHODS

PDB

The Protein Data Bank (August 26, 1998 version, containing 6247 entries) was searched for thiocyanate and azide containing proteins using the 3DB Browser. The modes of searching by keyword and text queries "thiocyanate" and "azide" in the complete PDB text and the additional search with associated group "thiocyanate" and "azide" or SCN and N3 were used. A summary of the currently available protein crystallographic structures with anions is provided in Table I (thiocyanate) and Table II (azide). The complete PDB file with crystallographic coordinates for each protein structure was retrieved and anion binding site was inspected by SYBYL, RasMol and Turbo-Frodo [34] computer graphics packages. The anion...metal and anion...hydrogen donor group interactions were analysed for all crystallographically independent parts of proteins with the criteria of M/D...N(anion) ≤ 3.6 Å, M/D...S(anion) ≤ 4.0 Å.

CSD

Related structures were retrieved from the CSD, Version 5.12. Substructure search, geometry calculation and data analysis were performed *via* the programs *QUEST3D* and *VISTA* (CSD) for three different groups of compounds: (I) ionic crystals containing discrete anion; (II) metal...anion complexes and (III) Element...anion. The metal M was restricted to groups 1A, IIA, IIIA, IB, IIB, VIB, VIIB and VIII metals. The element EE was restricted to boron, carbon, silicon, VA, VIA and VIIA groups of elements. Structures were accepted without any R-factor screening, but were required to have error-free coordinate set, to run successfully through each of the CSD check procedures, and to exhibit no crystallographic disorder. Polymeric complexes were excluded.

The following geometrical parameter were calculated for thiocyanate: d1(N-C), d2(S-C) and ∠(SCN); for azide: d1(N1-N2), d2(N2-N3) and ∠(NNN). The Metal...Anion interaction was characterized by the additional parameters: d(M...A) and ∠(MAC) for SCN and ∠(MNN) for N3. Limiting contact criteria of M...N < 3.6 and M..S ≤ 4.0 Å were used. Hydrogen Donor...Anion interaction were analysed for each group of compounds. The D-H distance (D=O, N) was normalised using geometrical criteria [35]. The following values were used as non-bonded contacts criteria: D...N≤ 3.6, D...S ≤ 4.0 Å and 90≤ ∠(DHA) ≤ 180°.

Acknowledgements

The author is grateful to Dr. Alex Yanovsky for helpful suggestions and discussion. Dr. C. Pascard is thanked for her most needed help at the beginning of this study.

References

1. *Supramolecular Chemistry of Anions*, Bianchi, A., Bowman-James, K. and García España, E. (Eds), Wiley-VCH, 1997; *Inclusion Chemistry*, Atwood, J.L., Davies, J.E.D. and MacNicol, D.D. (Eds), Oxford University Press, Oxford, 1991.
2. Chakrabarti, P. (1990). *Biochemistry*, **29**, p. 651; Schmidbaur, H., Classen, H.G. and Helbig, J. (1990). *Angew. Chem. Engl.*, **29**, p. 1090; Glusker, J.P. *Structural Aspects of Metal Liganding to Functional Groups in Proteins, in* "Advances in Protein Chemistry" (Anfinsen, C.B., Edsall, J.T., Richards F.M. and Eisenberg, D.S. (Eds), Vol **42**. *Metalloproteins: Structural Aspects*. Academic Press, Inc. 1991, p. 3; Alexander, R.S., Kanyo, Z.F., Chirlian, L.E. and Christianson, D.W. (1990). *J. Am. Chem.* **112**, p. 933; Görbitz, C.H. and Etter, M.C. (1992). *J. Chem. Soc. Pekin Trans. 2* (1). p. 131; Chakrabarti, P. (1993). *J. Mol. Biol.* **234**, p. 463; Pathaneni, S.S. and Desiraju, G.R. (1993). *J. Chem. Soc., Dalton Trans.*, p. 2505; Glusker, J.P. (1995). *Acta Cryst.* **D51**, p. 418; Schneider, B., Kabelac, M. and Hobza, P. (1996). *J. Am. Chem. Soc.* **118**, p. 12207.
3. Ducruix, A. and Giegé, R. *Crystallization of proteins and nucleic acids*, IRL/Oxford Press, 1992.
4. Allen, F.H., Bellard, S., Brice, M.D., Cartwright, B.A., Doubleday, A., Higgs, H., Hummelink, T., Hummelink-Peters, B.G., Kennard, O., Motherwell, W.D.S., Rodgers, J.R. and Watson, D.R. (1979). *Acta Cryst.*. B35, p. 2331.
5. Bernstein, F.C., Koetzle, T.F., Williams, G.J.B., Meyer, E.F. Jr., Brice, M.D., Rodgers, J.R., Kennard, O., Shimanouchi, T. and Tasumi, M. (1977). *J. Mol. Biol.* **112**, p. 535.

6. *Advanced Inorganic Chemistry*, Cotton, F.A. and Wilkinson, G. (Eds), John Willey & Sons, 1988.
7. Howell, P.L. (1995). To be published. Deposited as PDB code 1TEW.
8. Saludjian P., Prange, T., Navaza, J., Ménez, R., Guilloteau J.-P., Riès-Kautt, M. and Ducruix, A. (1992). *Acta Cryst.* **B48**, p. 520.
9. Whittingham, J.L., Chauhuri, S., Dodson, E.J., Moody, P.C.E. and Dodson, G.G. (1995). To be published, Deposite as PDB 2TCI.
10. Eriksson, A.E., Kylsten, P.M., Jones, T.A. and Liljas, A. (1988). *Proteins. Struct., Funct., Genet.*, **4**, p. 283.
11. Conti, E., Moser, C., Rizzi, M., Mattevi, A., Lionetti, C., Coda, A., Ascenzi, P., Brunori, M. and Bolognesi, M. (1993). To be published. Deposite as PDB code 2FAM.
12. Lan, M.S., Dixon, M., Pattridge, K.A., Stallings, W.C., Fee, J.A. and Ludwig, M.L. (1994). To be published. Deposited as PDB code1ISC; Lan, M.S., Dixon, M., Pattridge, K.A., Stallings, W.C., Fee, J.A. and Ludwig, M.L. (1994). To be published. Deposited as PDB code 1MNG.
13. Messerschmidt, A. and Wever. R. (1995). To be published. Deposited as PDB code 1VNC.
14. Messerschmidt, A., Luecke, H. and Huber, R. (1993). *J. Mol. Biol.* 220, p. 994.
15. Allendorf, M.D., Spira, D.J. and Solomon, E.I. (1985). *Proc. Natl. Acad. Sci. U.S.A.*, **82**, p. 3063.
16. Lamzin, V.S., Dauter, Z., Popov, V.O., Harutyunyan, E.H. and Willson, K.S. (1994). *J. Mol. Biol.* **236**, p. 759.
17. Krebs, J.F., Ippolito, J.A., Christianson, D.W. and Fierke, C.A. (1993). To be published. Deposited as PDB code 1CVA.; Jonsson, B.M., Hakansson, K. and Liljas, A. (1993). FEBS. Lett.. **322**, p. 186; Scolnick, L.R. and Christianson, D.W. (1996). To be published. Deposited as PDB code 1UGB. Scolnick, L.R. and Christianson, D.W. (1996). *Biochem.*. **35**, p. 16429.
18. Smith, J.L., Hendrickson, W.A. and Addison, A.W. (1983). *Nature*, **303**, p. 86; Holms, M.A. and Stenkamp, R.E. (1991). *J. Mol. Biol.* **220**, p. 723; Sheriff, S., Hendrickson, W. A. and Smith, J.L. (1987). *J. Mol. Biol.* **197**, p. 273.
19. Rizzi, M., Ascenzi, P., Coda, A, Brunori, M. and Bolognesi, M. (1992). To be published. Deposited as PDB code 1SWM; Mattevi, A., Gatti, G., Coda, A., Ascenzi, P. and Brunori, M. (1991). *J. Mol. Recog.* **4**, p. 1.
20. Tsai, L.-C., Bonander, N., Harata, K. Karlsson, B.G., Vanngard, T., Langer, V. and Sjolin, L. (1996). To be published. Deposited as PDB code 2TSB.
21. Kiiefer, L.L., Paterno, S.A. and Fierke, C.A. (1995). *J. Am. Chem. Soc.*, **117**, p. 6831.
22. Tchertanov, L. and Pascard, C. (1997). Acta Cryst.. **B53**, p. 904; Tchertanov, L. and Pascard, C. *Molecular recognition of anionic species: hydrogen-bonding properties of sulfate and thiocyanate* in Molecular Recognition and Inclusion,. Coleman, A.W. (Ed), 1998, Kluwer Academic Publishers p. 523.
23. Pauling, L. *The Nature of the Chemical Bond*, 3rd ed. Ithaka: Cornell University Press, 1960; Wells, A.F. *Structural Inorganic Chemistry*. Fifth edition; Clarendon Press. Oxford, 1984.
24. Pearson R.G. (1963). *J. Am. Chem. Soc.* **85**, p. 3533; Ahrland, S. Chatt, J. and Davies, N. R. (1958). *Q. Rev. Chem. Soc.* **12**, p. 265; Jain, P.C. and Lingafelter, E.C. (1967). *J. Am. Chem. Soc.* **89**, p. 6131; Cannas, M., Carta, G., Cristini, A. and Marongiu, G. (1977). *Inorg. Chem.*. **16**, p. 228.
25. Musaev, D.G., Yakobson, V.V. and Charkin, O.P. (1989). *Koordin. Khim.* **15**, p. 1011; Musaev, D.G., Makhaev, V.D. and Charkin, O.P. (1991). *Koordin. Khim.* **17**, p. 548; Parrini, F., Morales, R.G.E. (1993). *J. Mol. Struc. (Theochem)* **282**, p. 59.
26. Treinin, A. in *The Chemistry of the Azido Group* (Ed. S. Patai), John Wiley & Sons, London (1971), p. 1; Kaftory, M. in *The Chemistry of halides, pseudo-halides and azides*. Part 2. Patai, S. and Rappoport, Z. (Eds)., John Wiley and Sons, London (1983), p. 1229.
27. Klapötke, Th., M., (1997). *Chem. Ber./Recueil*, 130, p. 443.
28. Allen, F.H. and Garner, S.E. *Structural chemistry of triple-bonded groups* in *Supplement* C2: Edited by S. Patai © John Wiley @ Sons Ltd. 1994.
29. Schleyer, P.R. and Kos, A.J. (1983). *Tetrahedron*, **38**, 7 p. 1141.
30. Tchertanov, L. and Pascard, C. (1996). *Acta Cryst.*. **B52**, p. 685.
31. Tchertanov, L. (1999).). *Acta Cryst.* B55, p. 807.
32. Pearson, R.G. (1963). *J. Am. Chem.* Soc. **85**, N°. 22, p. 3533.
33. Xu, Y., Sellam, O., Moréra, S., Sarfati, S., Biondi, R., Véron, M. and Janin, J. (1997). *Proc. Natl. Acad. Sci. USA.* **94**, p. 7162.
34. SYBYL Molecular Modeling System (1991). Version 6.3, Tripos Ass., 1699 Handley Road, St. Louis, MO; RasMol Molecular Visualisation Program (1994). Version 2.5. A, Sayle, R. Biomolecular Structure Glaxo Research and Development, Greenford, Middlesex, UK; Turbo-Frodo 3-D Molecular Modelling Graphics Software (1997). Version 5.4.
35. Taylor, R. and Kennard, O. (1984). *Acc. Chem. Res.* **17**, p. 320.