First Crystal Structure for a Gold Carbene–Protein Adduct

Giariita Ferraro,† Chiara Gabbiani,‡ and Antonello Merlino*,†

†Department of Chemical Sciences, University of Naples Federico II, Complesso Universitario di Monte Sant’Angelo, Via Cintia, I-80126 Napoli, Italy
‡Department of Chemistry and Industrial Chemistry, University of Pisa, Via Moruzzi 13, 56124 Pisa, Italy

Supporting Information

ABSTRACT: The X-ray structure of the adduct formed in the reaction between the gold N-heterocyclic carbene compound Au(NHC)Cl (with NHC = 1-butyl-3-methyl-imidazole-2-ylidene) and the model protein thaumatin is reported here. The structure reveals binding of Au(NHC)+ fragments to distinct protein sites. Notably, binding of the gold compound occurs at lysine side chains and at the N-terminal tail; the metal binds the protein after releasing Cl– ligand, but retaining NHC fragment.

N-heterocyclic carbene gold(I) compounds [Au(I)(NHC)R], where R is an auxiliary ligand (for instance, −Cl, −NHC, or −PPh3), are linear species,1−3 either neutral or cationic, widely used in homogeneous catalysis4−10 and in medicinal chemistry.11−14 Thanks to their favorable and peculiar chemical and biological properties, gold(I) N-heterocyclic carbene compounds have been tested extensively in recent years as anti-arthritic,15 antibacterial,16 and anticancer agents.17−26

The anticancer activity of these compounds seems to arise from potent and selective inhibition of the enzyme thioredoxin reductase (TrxR), particularly in cancer cells,27−38 though this mechanistic issue is still highly controversial and debated. Electrospray ionization mass spectrometry (ESI MS) measurements suggest that gold carbene compounds preferentially bind to free cysteine (or seleno-cysteine) side chains,39,40 with AuNHC+ fragments or Au+ ions coordinating the thiol (selenol) group upon release of the carbene ligand(s).39,40 Although the interaction of gold carbene compounds with many proteins (for example, TrxR, phosphatases, glutathione reductases, serum albumin, Atox-1) has now been studied,19,27,43−46 structural data on the adducts formed upon reaction still lack. Here, we report the first X-ray structure of the complex formed in the reaction between a protein, i.e., thaumatin, and the gold carbene compound Au(NHC)Cl (Compound 1, Figure 1), where NHC is 1-butyl-3-methyl-imidazole-2-ylidene. Thaumatin was chosen as a model system since it is frequently used as the archetypal protein for crystallization studies.34−46

Crystals of the thaumatin/Au(NHC)Cl adduct were obtained by the soaking procedure starting from thaumatin crystals grown by the hanging drop vapor diffusion method using a reservoir solution of 0.2 M sodium tartrate pH 7.2 and 20% polyethylene glycol (see Supporting Information for details). X-ray diffraction data were collected on these crystals at 1.70 Å resolution (Table S1).

The overall structure of the Thaumatin/Au(NHC)Cl complex is shown in Figure 2. The structure, containing 2020 non hydrogen atoms, refines to a R-factor of 0.170 (R-free = 0.187). The overall fold of the protein is hardly affected by the gold carbene binding. Carbon alpha root-mean-square deviation from the native protein (PDB code 3QY5)47 is 0.10 Å. Six gold compound fragments have been identified bound to thaumatin. Binding of the gold compound occurs at lysine side chains and at the N-terminal tail; the metal binds the protein after releasing Cl– ligand, but retaining NHC fragment.

Figure 1. Structure of compound 1, the gold carbene complex used in this study.
Notably, in our structure, at least in three distinct binding sites, i.e., close to N-terminal amine (Figure 3A), Lys47 and Lys106 side chains (Figure 3B), the NHC ligand is retained upon protein adduct formation. In these sites, N–Au–C bond angles were in the range 177.2–178.9°; N–Au distances were on average equal to 2.0 Å and all the Au–C carbene bonds distances were in the range 1.9–2.0 Å. Metalated side chains of Lys106 adopt two distinct conformations (Figure 3B). All the residues involved in the Au(NHC)Cl recognition are solvent exposed (solvent accessible surface areas calculated using native protein are in all cases >22 Å²).

According to occupancy factors, N-terminus and side chains of Lys49 are the sites with higher affinity for gold fragments. Lower affinity is observed in the case of Lys97 and Arg175. In the binding site close to Lys97, where the Au occupancy is very limited (=0.15), the electron density is not sufficiently well-defined to identify the other metal ligand (Figure S1). Similarly, in the binding site close to Arg8, a gold ion is present. In this site there is not enough space to accommodate the NHC ligand (Figure S2). Close to Arg175, a gold ion binds a ligand, which is probably a water molecule. The possibility that this ligand could be a Cl⁻ ion is excluded by the inspection of the anomalous electron density map. However, it should be noted that in this site the interpretation of the map is complicated by the presence of alternative conformations of the Arg side chain (Figure S3).

A summary of the structural features of the Au(NHC)Cl binding sites and of the interactions that the gold compound fragments form with surrounding protein residues are reported in Table S2. Notably, our results have been confirmed by the inspection of the electron density maps obtained using an additional data set collected, at 1.93 Å resolution, on a second Thaumatin/Au(NHC)Cl crystal (Figure S4).

The finding that NHC is kept close to Au atoms upon protein binding is quite remarkable; previous ESI MS data carried out on the adduct formed in the reaction of the same compound with Atox-1 revealed a release of NHC moiety from gold center.18,19,39,40 According to hard–soft acid–base theory,48 Au(I) is considered a soft metal ion; therefore, it should exhibit a marked preference for soft ligands such as thiols of free cysteines and methionines.49 However, it was previously demonstrated by X-ray crystallography that Au(I) ions can coordinate solvent exposed His,50–53 also in the presence of free thiols. It has also been shown that Au(I) can bind also Gln side chains, although with lower affinity than that exhibited for Cys, Met and His residues.54 Our structural data demonstrate that in the absence of free cysteines and histidines, these gold compounds are able to bind N-terminal amino group, the side chains of Lys or even Arg residues.

In conclusion, here we have reported the first crystal structure of an adduct formed between a gold(I) carbene compound and a protein. Although our data have been obtained using a model protein and not the real pharmacological target of these compounds, they reveal a feature that could have general significance: N-terminus, Lys, and Arg side chains may act as targets for these compounds; NHC moiety may be retained upon protein binding.
The authors thank G. Sorrentino and M. Amendola for technical assistance. C.G. acknowledges Beneficentia Stiftung (Vaduz, Liechtenstein) and AIRC (IG-12085) for generous financial support.

ACKNOWLEDGMENTS
The authors declare no competing financial interest.

REFERENCES


