Sulphur-aromatic interactions in proteins

K.S.C. Reid*, P.F. Lindley and J.M. Thornton†

Crystallography Department, Birkbeck College, University of London, Malet Street, London, WC1E 7XH, England

Received 9 August 1985

The geometry of sulphur-aromatic interactions in globular proteins has been analysed using crystallographic data derived from 36 proteins, solved to resolutions of 2 Å or better. About half of all sulphur atoms from cyst(e)ine and methionine residues are in contact (≦6 Å from ring centroid) with an aromatic ring (phenylalanine, tyrosine or tryptophan). Compared to carbon and nitrogen atoms the interacting sulphur atoms express an affinity towards the edge of the aromatic rings, and avoid the region above the ring in the vicinity of the π-electrons. This preference is similar to that previously found for oxygen atoms around phenylalanine rings, and may be electrostatic in origin.

Sulfur-aromatic interaction *Protein* *Side-chain contact* *Cyst(e)ine* *Methionine* *Packing*

1. INTRODUCTION

It has been suggested that a sulphur atom can interact favourably with an aromatic ring, via an S-π interaction [1,2]. Sulphur-aromatic interactions are commonly observed in the hydrophobic core of proteins and may contribute to protein stability. They are particularly common in the eye lens protein γ-crystallin [3] where they may help to protect the protein against radiation damage [4]. In this paper we investigate the frequency and geometry of such interactions by analysis of coordinate data of high resolution protein structures. Interactions between the sulphur atoms of cysteine and methionine and the aromatic rings of phenylalanine, tyrosine and tryptophan are considered.

2. METHODS

The analysis of sulphur-aromatic (herein after S-aromatic) interactions involved searching through available structural data for sulphur atoms in the contact vicinity of aromatic rings.

2.1. *Co-ordinate data*

These were obtained from the Cambridge Protein Data Bank [5]. Only non-homologous structures obtained to nominal resolution ≦ 2.5 Å were included. The 36 proteins with their data bank filenames are given: avian pancreatic polypeptide, 1PPP; L-arabinose-binding protein, 1ABP; actinidin, 2ACT; liver alcohol dehydrogenase, 4ADH; phospholipase A_1, 1BP2; cytochrome b_5, 2B5C; carbonic anhydrase C, 1CAC; concavalin A, 2CNA; carboxypeptidase A, 1CPA; parvalbumin, 3CPV; crambin, 1CRN; cytochrome *c*, 4CYT; cytochrome C551, 251C; erythrocruorin, 1ECD; human immunoglobulin FABL fragment, 2FAB; ferredoxin, 1FDX; ferredoxin, 1FD1; flavodoxin, 4FXN; high potential iron protein, 1HIP; lactate dehydrogenase, 4LDH; leg haemoglobin, 1HML; lysozyme (hen egg-white), 2LYZ; lysozyme (phage T_4), 1LZM; neurotoxin, 1NXB; prealbumin, 2PAB; plastocyanin, 1PCY; pancreatic trypsin inhibitor, 3PTI; trypsin, 1PTN; Bence-Jones REI protein, 1REI; Rhodanese, 1RHD; ribonuclease A, RNS; rubredoxin, 3RXN; subtilisin, 1SBT; superoxide dismutase, 2SOD; triose phosphate isomerase, 1TIM; thermolysin, 2TLN.

* Present address: Blackett Laboratory, Imperial College of Science and Technology, Prince Consort Road, London SW7 2BZ, England

† To whom correspondence should be addressed

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/85/$3.30

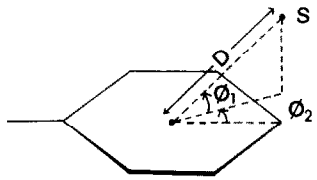

Fig.1. Parameters used to define relative orientation. $|D|$, distance between sulphur atom and ring centroid. ϕ_1, angle of elevation between vector $|D|$ and the ring plane. ϕ_2, equatorial angle defined by the projection of D into the ring plane and a vector through one of the ring atoms; Cz for Phe and Tyr, Cz and NE_1 for the 6- and 5-membered rings of Trp.

2.2. *Definition of parameters describing relative orientation*

Three parameters, $|D|$, ϕ_1 and ϕ_2, were used to define the orientation of sulphur atoms relative to aromatic rings (fig.1).

($|D|$) The distance between the sulphur atom and the centroid of the aromatic ring. The indole ring of tryptophan is treated as two separate components – a 6 and 5 membered ring.

(ϕ_1) Elevation angle. This is an index of co-planarity between the sulphur atom and the ring plane, and the values range from 0 (in-plane) to 90° (directly above, or perpendicular to, the ring plane).

(ϕ_2) Equatorial angle. This is the point on the compass (described by the ring-plane) from which the sulphur atom approaches relative to a meridian, itself fixed by one of the ring atoms, and the values range from 0 to 180° for Phe and Tyr and from 0 to 360° for Trp.

The ranges for ϕ_1 and ϕ_2 do not encompass all space around the ring centroid, but are sufficient to define a unique relative orientation of a neighbour atom because of the mirror-symmetry inherent in aromatic side-chains.

2.3. *Normalisation*

The volumes defined by equiangular intervals of ϕ_1 are not uniform but are proportional to $\sin(\phi_1^U - \phi_1^L)$, where the superscripts U and L refer to the upper and the lower angular bounds of the interval respectively. However aromatic ring hydrogens occlude volume in the low ϕ_1 region, therefore the observed distribution of all atoms in ϕ_1 space provides a control which bears closer

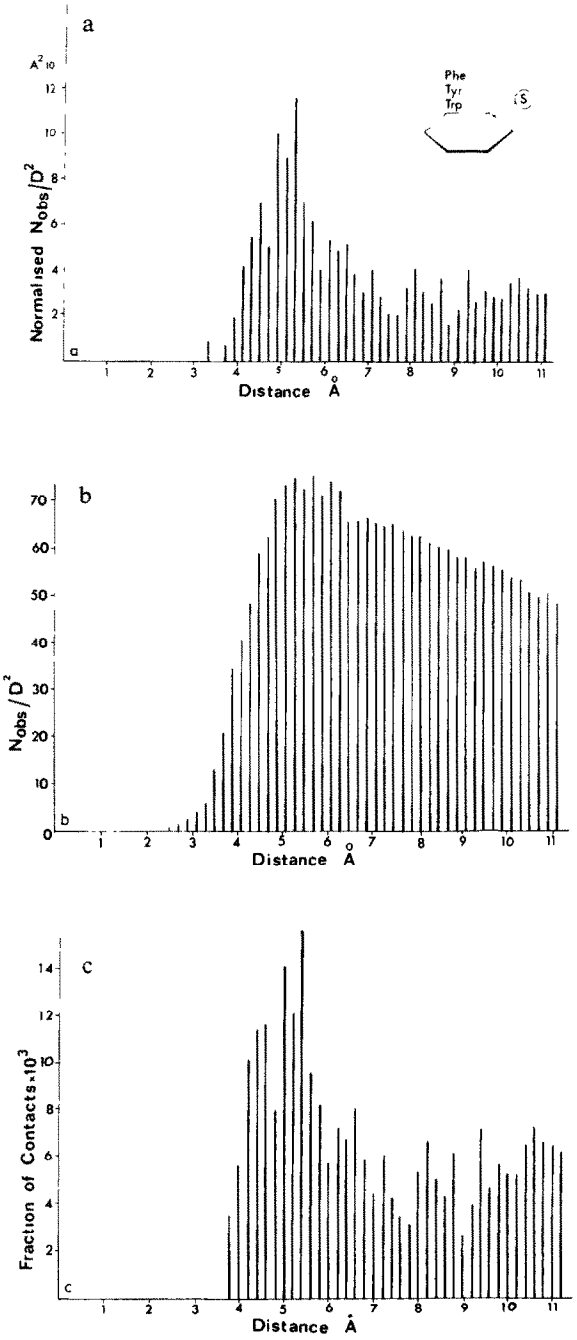

Fig.2. Radial distribution plots for atoms about aromatic rings (Phe, Tyr and Trp) taken from 36 proteins. (a) Distribution of D values for all-sulphur (Cys and Met) about all aromatic rings. (b) Distribution of D values for all atoms (C,N,O and S) about all aromatic rings. (c) Normalised D distribution for all sulphur atoms, expressed as a percentage of 'all-atom' contacts.

parity with the actual volume available. The distribution frequencies are expressed as percentages of the total number of atomic contacts.

3. RESULTS

3.1. *Radial distance distribution*

The parameter $|D|$ was calculated for every sulphur atom and every aromatic group in each protein, for all of the 36 proteins. The radial distribution plots of all-sulphur all-aromatic $|D|$ values (fig.2a) rise to a sharp peak at $|D| \simeq 5.3$ Å. Comparison with the corresponding plots for the 'all-atom' radial distribution (fig.2b), which demonstrate a much broader peak and do not fall off as sharply, suggests that sulphur atoms prefer to be in van der Waals' contact with aromatic rings. The distribution of sulphur atoms expressed as a percentage of 'all-atom' contacts (fig.2c) emphasizes this deviation from the norm, most marked at $|D| \simeq 5.3$ Å.

3.2. *Frequency of sulphur-aromatic contacts*

For the purpose of this paper an S-aromatic interaction is operationally defined as

$$\{|D| \leqq 6 \text{ Å}\}$$

Table 1

Sulphur-aromatic interactions*

	Phe	Tyr	Trp	Cysteine	Cystine	Met
Total no. of residues	225	210	110	67	2×39 = 78	90
Residues which participate in S-aromatic*	55	38	27	29	40	51
% of interacting residues	24	18	25	43	51	57
No. of interactions provided by residue types	78	50	37	40	56	86
No. of interactions per interacting residue	1.4	1.3	1.4	1.4	1.4	1.7

*For interaction $\{|D| \leqq 6.0$ Å$\}$

This range will include all sulphur atoms in van der Waals' contact with an aromatic ring. In this way 185 S-aromatic interactions were found, analysis of which is given in table 1. Almost half the sulphydryl cysteine and more than half the cystine and methionine sulphurs are within 6 Å, or interact with an aromatic ring (cf. only ⅓ of Arg-NH_2^+ participate in salt bridges [6]). In contrast not more than ¼ of the aromatic groups are found to interact with sulphur atoms, the proportion of interacting vs non-interacting tyrosine being less than that of phenylalanine and of tryptophan. Of those interacting residues, many are found to be associated with more than one partner: 51 methionines provided 86 interactions, suggesting that almost 70% of the interacting methionines have two aromatic residues in contact. The results of Warme and Morgan [7,8] also showed that sulphur-aromatic carbon contacts occur slightly more frequently than expected.

3.3. ϕ_1 *distribution*

The ϕ_1 distribution for all S-aromatic interactions is shown in fig.3a. For comparison the distribution of sulphur, carbon, oxygen and nitrogen atoms about phenylalanine rings is shown in fig.3b. As expected from volume considerations (dashed curve, fig.3a) the frequencies approximately fall off with increasing ϕ_1. There are very few interactions where the sulphur approaches the ring face. However there are distinct differences for the different atom types. The nitrogen and carbon atom plots peak at the 20–30° and 40–50° intervals, respectively, whilst oxygen and sulphur atoms, which show very similar distributions, achieve their maximum value in the 0–10° interval. Thomas et al. [9] have shown that oxygen preferentially approaches aromatic rings in-plane, and normalised plot of the S-phenylalanine ϕ_1 distribution (fig.3c) shows a similar linear relationship to the one they obtained for oxygen-phenylalanine. This suggests that interacting sulphur expresses an affinity towards the edge of the aromatic rings and, compared to carbon and nitrogen atoms, avoids the region above the ring in the vicinity of the π-electrons.

3.4. ϕ_2 *distribution*

The distribution of ϕ_2 values for sulphur atoms (not shown) is comparable with that obtained for

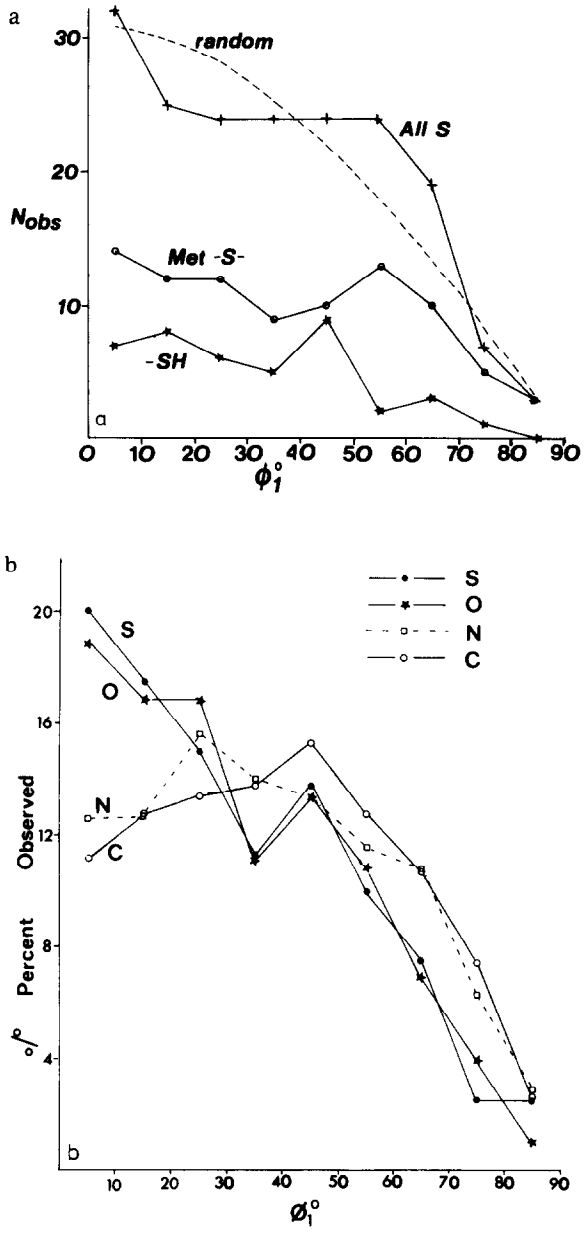

Fig.3. ϕ_1 distribution for atoms at distances $\leqq 6$ Å from ring centroids. (a) ϕ_1 distribution of all sulphur atoms about all aromatic rings. The dashed line shows the 'random' distribution derived from volume considerations. Plots were also made for the different 'types' of sulphur atoms (ie. methionine, cysteine and cystine). All the plots were broadly similar, although the number of observations is low. (b) ϕ_1 distribution for C,N,O and S atoms about phenylalanine rings. For each atom type, the percentage of atoms of that type in the specified angular range is plotted. (c) Normalised ϕ_1 S-phenylalanine distribution expressed as a proportion of total contacts. The straight line is the least-squares fit, suggesting that an edge-on interaction (0°) is approx. 4-times as 'favourable' as the ring face interaction. The S-tyrosine distribution (not shown) is similar.

other atom types, showing a broad distribution over all angles, excepting the 180° region, which is forbidden by the CB atom. In addition for tyrosine the low ϕ_2 values are underpopulated due to exclusion by the -OH group.

4. DISCUSSION

The aromatic and sulphur-containing amino acids are non-polar and, as a consequence, occur most frequently in the interior of proteins. This accounts for the observed frequency of interaction being larger than might be expected for random association [7,8]. However from the geometry of such interactions it appears that sulphur atoms, unlike carbon and nitrogen, predominantly approach the edge of aromatic rings rather than interact in a planar stacking fashion. The use of the

term S-π interaction in this context is therefore misleading [1,2].

A similar mode of interaction is demonstrated by oxygen-phenylalanine systems and the quantum mechanical calculations of Thomas et al. [9] suggest that the origins of this effect are electrostatic. The negatively charged oxygen is attracted by the positively charged aromatic ring hydrogens, and repelled by the electron-rich π-cloud. It is reasonable to propose a similar rationale for the results involving sulphur atoms since sulphur is electronegative. However, the work of Momany et al. [10] ascribes positive partial charges of 0.015, 0.010, and 0.035 to cysteine, cystine and methionine sulphur, respectively. The linear distribution shown in fig.3c reflects not only the sulphur distribution, but also the distributions for all the other atoms, which may themselves exhibit energy-defined preferences. A complete quantum mechanical calculation is needed to understand the origins of the observed distribution. However, the results of this paper demonstrate that even in the protein core, where the conformation is constrained to fulfil the requirements of close packing and hydrogen bonding, there are distinct differences between the spatial distributions about aromatic rings for the different atom types.

REFERENCES

[1] Morgan, R.S. and McAdon, J.M. (1980) Int. J. Peptide Prot. Res. 15, 177–180.
[2] Morgan, R.S., Tatsch, C.E., Gushard, R.H., McAdon, J.M. and Warme, P.K. (1978) Int. J. Peptide Prot. Res. 11, 209–217.
[3] Blundell, T.L., Lindley, P.F., Miller, L.R., Moss, D.S., Slingsby, C., Tickle, I.J., Turnell,W.G. and Wistow, G.J. (1981) Nature 289, 771–777.
[4] Summers, L., Wistow, G., Narebor, M., Moss, D., Lindley, P., Slingsby, C., Blundell, T.L., Bartunik, H. and Bartels, K. (1984) Peptide and Protein Reviews 3, 147–168.
[6] Barlow, D.J. and Thornton, J.M. (198) J. Mol. Biol. 168, 867–885.
[7] Warme, P.K. and Morgan, R.S. (1978) J. Mol. Biol. 118, 273–302.
[8] Warme, P.K. and Morgan, R.S. (1978) J. Mol. Biol. 118, 289–304.
[9] Thomas, K.A., Smith, G.M., Thomas, T.B. and Feldmann, R.J. (1982) Proc. Natl. Acad. Sci. USA 79, 4843–4847.
[10] Momany, F.A., McGuire, R.F., Burgess, A.W. and Scheraga, H.A. (1975) J. Phys. Chem. 79, 2361–2381.