doi:10.1016/j.jmb.2008.10.086 J. Mol. Biol. (2009) 385, 595–605

JMB

Available online at www.sciencedirect.com

ScienceDirect

ELSEVIER

Halogenated Benzenes Bound within a Non-polar Cavity in T4 Lysozyme Provide Examples of I···S and I···Se Halogen-bonding

Lijun Liu, Walter A. Baase and Brian W. Matthews*

Institute of Molecular Biology, Howard Hughes Medical Institute, University of Oregon, Eugene, OR 97403, USA

Department of Physics, University of Oregon, Eugene, OR 97403, USA

Received 5 September 2008; received in revised form 23 October 2008; accepted 29 October 2008
Available online
6 November 2008

We showed earlier that the mutation of Leu99 to alanine in bacteriophage T4 lysozyme creates an internal cavity of volume ~150 Å^3 that binds benzene and a variety of other ligands. As such, this cavity provides an excellent target to study protein-ligand interaction. Here, we use low-temperature crystallography and related techniques to analyze the binding of halogen-incorporated benzenes typified by C_6F_5X, where X=H, F, Cl, Br or I, and C_6H_5X, where X=H or I was also studied. Because of the increased electron density of fluorine relative to hydrogen, the geometry of binding of the fluoro compounds can often be determined more precisely than their hydrogen-containing analogs. All of the ligands bind in essentially the same plane but the center of the phenyl ring can translate by up to 1.2 Å. In no case does the ligand rotate freely within the cavity. The walls of the cavity consist predominantly of hydrocarbon atoms, and in several cases it appears that van der Waals interactions define the geometry of binding. In comparing the smallest with the largest ligand, the cavity volume increases from 181 Å^3 to 245 Å^3. This shows that the protein is flexible and adapts to the size and shape of the ligand. There is a remarkably close contact of 3.0 Å between the iodine atom on C_6F_5I and the sulfur or selenium atom of Met or SeMet102. This interaction is 1.0 Å less than the sum of the van der Waals radii and is a clear example of a so-called halogen bond. Notwithstanding this close approach, the increase in binding energy for the halogen bond relative to a van der Waals contact is estimated to be only about 0.5–0.7 kcal/mol.

Edited by R. Huber

Keywords: cavity; iodine–sulfur interaction; iodine–selenium interaction

Introduction

Molecular recognition is key to understanding various biological processes such as enzyme–substrate recognition, receptor–ligand interaction, antigen-antibody binding, and protein–nucleic acid interactions, among others. Molecular docking, aimed at predicting the affinity and binding mode of a ligand to a protein target or receptor has had some success but still needs further improvement. In particular, most docking algorithms ignore protein flexibility even though increasing evidence suggests that the binding sites of proteins can adapt to the presence of the ligand. For this reason, among others, docking a small molecule ligand is still a huge challenge.

Within this laboratory, the mutation of Leu99 to Ala in bacteriophage T4 lysozyme (L99A) has been shown to produce an internal non-polar cavity[1] that binds benzene as well as a variety of apolar derivatives.[2-4] The effect of introducing a polar amino acid into the wall of the cavity has also been examined.[5] Also, the noble gases, argon, krypton and xenon, have been used to explore several cavity-containing T4 lysozyme mutants including L99A.[6] In the present study, we used benzene-halogen derivatives (Table 1) to explore the polarizability and flexibility of the L99A cavity, and to examine halogen–protein interactions.

*Corresponding author. E-mail address: brian@uoregon.edu.

Results

Screening for the halogen-incorporated benzene ligands

To explore the polarizability and tolerance of the L99A cavity in T4 lysozyme, we determined the crystal structures of the protein complexed with halogen-incorporated benzene derivatives (Table 1). The experimental procedures are described in Materials and Methods. All X-ray data were collected at 100K. In the first set of compounds, five positions on the phenyl ring were replaced with fluorine, while the sixth site varied from H, F, Cl, Br, to I. Replacement of hydrogen with fluorine does not change the molecular size or polarizability significantly (Table 2) but allows more precise determination of the ring geometry due to the increased electron density. To further examine the binding behavior of the L99A cavity, the fluoro-benzene derivatives with two or more chloride, bromide or iodide atoms were also examined, among others (Table 1).

Binding modes of C_6F_5X (X = H, F, Cl, Br, I) in the cavity of L99A

Benzene (C_6H_6)

C_6H_6 binds within the L99A cavity at 100 K (Fig. 1a) in essentially the same manner as at room temperature. Relative to the room temperature structure,[2] the rotation in the plane of phenyl ring is less than 7°, and the shift of the ring centroid is less than 0.15 Å. The C^{β} atom of Ala99 and the six carbon atoms of the phenyl ring form an almost regular hexagonal pyramid, with the six C^{β}–$C^{benzene}$ distances ranging from 3.6 Å to 3.9 Å. This defines the standard benzene binding site. In the walls of the cavity, the sulfur atom (S^{δ}) of the side chain of Met102 is the heaviest and most polarizable atom. Its closest approach to any of the benzene carbons is 4.3 Å.

Hexafluorobenzene (C_6F_6)

In an ($F_{obs,complex} - F_{obs,L99A}$) isomorphous difference electron-density map (Fig. 1b), the fluorine atoms of C_6F_6 are clearly visible and define the ring orientation and geometry. In contrast to benzene, the ring centroid of C_6F_6 undergoes a shift of 1.0 Å towards the side chain of Met102 , and there is also an out-of-plane rotation of about 21° (Fig. 1d; Table 2). As a result, two fluorine atoms from C_6F_6 approach S^{δ} of Met102 within 3.4 Å and 3.3 Å, respectively, which are typical van der Waals contacts. For convenience, this is denoted as the hexafluorobenzene binding mode. It does not occur for benzene itself.

Pentafluorobenzene (C_6HF_5)

The map showing the difference in density between C_6HF_5-L99A and L99A (Fig. 1c) has distinct density for four of the five fluorines, but weaker density at the two possible sites closest to Met102. Refinement suggests that the ligand adopts two different modes of binding with essentially equal occupancy. In one case the fifth fluorine atom is at the 11 o'clock position (Fig. 1c), 3.2 Å from S^{δ} of Met102. In the other mode (not shown), the fluorine is at the 9 o'clock site, 3.5 Å from the sulfur atom. Other than this 60° rotation, the other atoms retain essentially identical positions within 0.1 Å. Relative to benzene, the centroid of the C_6HF_5 ring shifts ~0.6 Å toward Met102 and there is an azimuthal rotation of about 14°.

The distance between the S^{δ} atom of Met102 and the hydrogen-occupied carbon atom of C_6HF_5 is 4.1 Å in either conformation. This may suggest a weak hydrogen bond (S···H-C) between S^{δ} and C–H from C_6HF_5.

Chloropentafluorobenzene (C_6ClF_5) and 1,3,5-trichloro-2,4,6-trifluorobenzene ($C_6Cl_3F_3$)

The difference electron density map (Fig. 2a) suggests that C_6ClF_5 binds in a mode similar to C_6F_6. The introduction of the heavier chlorine atom

Table 1. Crystallographic data for L99A at 100K and its complexes with halogen-incorporated benzenes

Ligand	Cell dimensions a = b (Å)	Cell dimensions c (Å)	Resolution (Å)	Total reflections	Unique reflections	Completeness (%)	R_{merge} (%)	R_{free} (%)	R_{work} (%)	PDB code
free	59.90	95.31	1.65	164,255	24,337	99.4	8.5	21.0	18.3	3DMV
C_6H_6	60.00	95.73	1.8	120,006	18,791	98.3	7.5	20.9	18.2	3DMX
C_6HF_5	59.76	94.86	1.8	168,278	18,531	98.6	7.2	22.3	19.4	3DN0
C_6F_6	60.06	95.48	2.0	140,224	13,862	99.3	13.0	21.4	17.8	3DMZ
C_6ClF_5	59.86	95.50	1.8	144,869	18,867	99.5	9.0	22.0	19.0	3DN1
C_6BrF_5	59.79	95.40	1.8	128,832	18,813	99.7	6.3	21.4	18.1	3DN2
C_6F_5I	59.95	96.18	1.8	127,099	17,811	93.0	4.7	21.0	18.1	3DN3
C_6H_5I	59.98	95.99	1.8	162,121	18,278	95.9	8.1	22.3	18.5	3DN4
$C_6Cl_3F_3$	59.77	95.34	1.8	118,028	18,504	98.2	5.2	22.1	18.4	3DN6
C_6F_5I[a]	60.03	95.96	1.7	181,614	22,609	99.6	5.7	22.3	19.7	3DN8
C_6H_5I[a]	60.38	96.77	1.7	202,764	23,070	99.8	7.6	21.7	18.3	3DNA

Compounds: C_6H_6, benzene; C_6HF_5, pentafluorobenzene; C_6F_6, hexafluorobenzene; C_6ClF_5, chloropentafluorobenzene; C_6BrF_5, bromopentafluorobenzene; C_6F_5I, iodopentafluorobenzene; C_6H_5I, iodobenzene; $C_6Cl_3F_3$, 1,3,5-trichloro-2,4,6-trifluorobenzene; $C_6Br_2F_4$, 1,4-dibromo-2,3,5,6-tetrafluorobenzene; $C_6F_4I_2$, 1,4-diiodo-2,3,5,6-tetrafluorobenzene.

[a] These values are for Se-L99A complexes.

Table 2. Geometry of ligand binding

Ligand	Polarizability of ligand[a] (Å^3)	Volume of ligand[b](Å^3)	Volume of cavity[c] (Å^3)	Shift of ligand[d] (Å)	Rotation of ligand[e] (°)
C_6H_6	10.4	81.2	181	–	–
C_6HF_5	9.6	111.5	229	0.6, 0.6	14,14
C_6F_6	9.6	117.6	233	1.0	21
C_6ClF_5	11.7	126.7	245	1.0	19
$C_6Cl_3F_3$	15.6	145.1	240	1.1	22
C_6BrF_5	12.9	130.8	224	0.5, 1.3, 1.7	8, 23, 18
$C_6Br_2F_4$	16.1	144.0	N/A	N/A	N/A
C_6H_5I	15.5	106.5	216	0.8, 1.6	0, 180
C_6F_5I	15.3	136.8	245	0.7	13
$C_6F_4I_2$	20.8	156.0	N/A	N/A	N/A

No binding was detected for the $C_6Br_2F_4$ and $C_6F_4I_2$.

[a] Molecular polarizability values were estimated from Miller & Savchik,[33] and Miller.[34]

[b] Volumes calculated from Zhao *et al*.[35]

[c] The cavity volumes were calculated with the ligand removed from the refined protein model. The experimental cavity volume for ligand-free L99A is 169 Å^3.

[d] The value quoted is the shift in the center of the benzene ring of the halogenated ligand relative to pure benzene. Where multiple values are quoted they correspond to different modes of binding.

[e] The ring of the ligand was superimposed on the ring of bound benzene by the combination of a rotation plus a translation along the axis of rotation. The value quoted is the rotational component. Multiple values correspond to different binding modes.

does not lead to distinct higher density at a single site, suggesting that this ligand utilizes a number of alternative modes. Refinements of the ligand with the chlorine in the six alternative positions gave occupancy values of the chlorine ranging from 0.15 in the mode shown in Fig. 2a, to 0.13 for all others. The

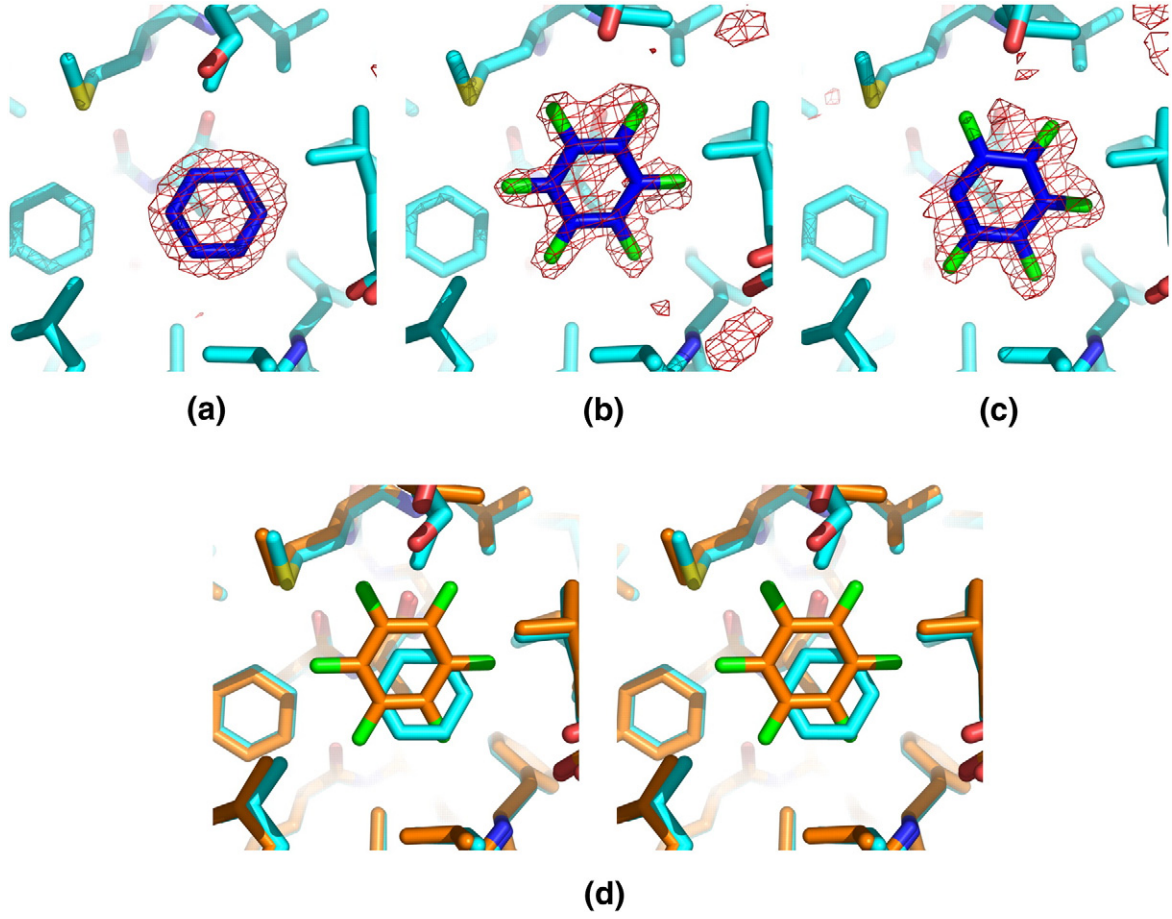

Fig. 1. Electron density maps showing the location of ligands bound in the L99A cavity. Maps have amplitudes ($F_{obs,complex} - F_{obs,L99A}$) and phases from the refined structure of L99A. Resolution as in Table 1; maps contoured at 3.0σ. (a) C_6H_6, ligand in dark blue. (b) C_6F_6, fluorine atoms in green. (c) C_6HF_5. An alternative binding mode, not shown, is rotated 60° clockwise. (d) Stereo figure superimposing the C_6F_6 complex (carbon atoms in orange, fluorine in green) on that for C_6H_6 (atoms in blue). The side chain of Met102 is at the top left with the sulfur atom in yellow. All figures were rendered with PyMOL [http://www.pymol.org].

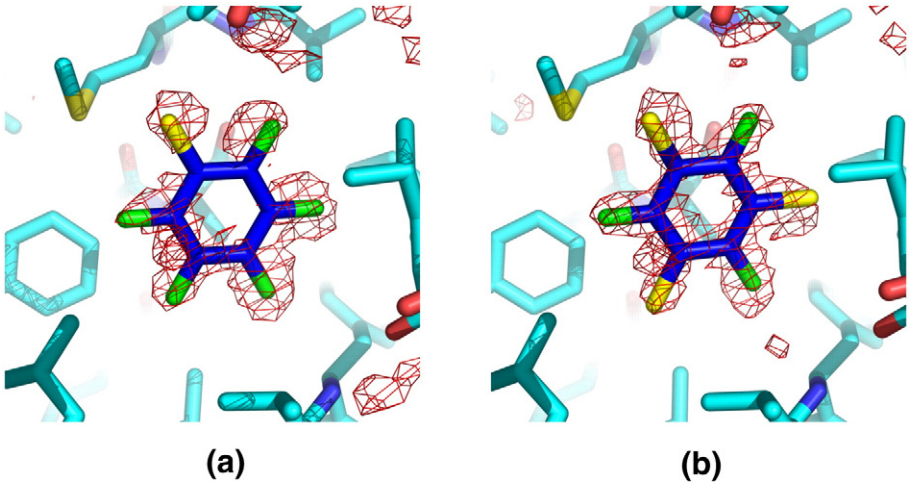

Fig. 2. Binding of C_6ClF_5 and $C_6Cl_3F_3$. (a) C_6ClF_5. Electron density map, calculated and shown as in Fig. 1a. The chlorine atom (in yellow) is disordered essentially equally between the location shown and the five fluorine sites. (b) Electron density for $C_6Cl_3F_3$. This ligand is also azimuthally disordered (see the text).

chlorine therefore appears to be distributed essentially equally over the six possible sites.

In the case of $C_6Cl_3F_3$ (Fig. 2b) it also appears that alternative modes of binding are adopted. The crystallographic refinement suggests that the conformation shown in Fig. 2b is slightly favored, with an occupancy of about 0.6. In the alternative mode, the ring is rotated by 60°. For the preferred modes of binding of both C_6ClF_5 and $C_6Cl_3F_3$, the distance from the closest chlorine atom to S^{δ} of Met102 is 3.4 Å. Among the four ligands, C_6HF_5, C_6F_6, C_6ClF_5 and $C_6Cl_3F_3$, the movements of the centroid of phenyl ring and changes in planar orientation relative to each other are less than 0.5 Å and 12°.

Iodopentafluorobenzene (C_6F_5I)

C_6F_5I binds to the L99A cavity in a single binding mode, as shown by the difference electron density (Fig. 3a). An anomalous density map (included in Fig. 3a) calculated with amplitudes ($F_+ - F_-$) from the C_6F_5I complex and with phases from the ligand-free L99A model has a very strong peak (38σ) at the iodine site inferred from the isomorphous difference map. Correspondingly, an anomalous peak of height 4.5σ could also be seen at the position of the S^{δ} atom of Met102. These peaks confirm the positions of the iodine and sulfur atoms. The iodine atom is located approximately colinearly between the carbon atom to which it is bound and the S^{δ} atom of Met102. The distance between I and S^{δ} is 3.0 Å, and the angle $S^{\delta}\cdots$I-C is 166°. The *B*-factors of the iodine and sulfur atoms are 22 $Å^2$ and 13 $Å^2$, respectively. There is a substantial change in the orientation of the ring. Relative to C_6H_6, the centroid of the phenyl ring is shifted by 0.5 Å (similar to C_6F_6) but there is an azimuthal rotation of approximately 30° around the axis perpendicular to the ring plane (Fig. 3c). Consequently, the van der Waals contacts between C_6F_5I and the surrounding atoms undergo a variety of changes. Included among these, the side chain of residue Leu118 rotates ~110° around its C^{β}-C^{γ} bond. We denote this conformation as the iodopentafluorobenzene binding mode.

Iodobenzene (C_6H_5I)

To further examine the I..S^{δ} interaction, we evaluated both C_6H_5I and 1,4-diiodotetrafluorobenzene ($C_6F_4I_2$). On the basis of a difference electron density map (not shown), $C_6F_4I_2$ is not detected to bind in the cavity, possibly because of its low degree of solubility and vapor pressure and/or larger size. In contrast, C_6H_5I does bind in the L99A cavity, but with two binding modes (Fig. 3b, d). The difference electron density is noisy but shows two strong peaks (both 7.0σ) 5.0 Å apart. An anomalous scattering map (Fig. 3b) also shows strong density (19σ and 15σ) at the two presumed iodine positions. The first binding conformation (not shown in Fig. 3b) is very similar to that of iodopentafluorobenzene with the iodine atom pointing at S^{δ} of Met102 (Fig. 3c). The distance I..S^{δ} is 3.3 Å, slightly longer than that for C_6F_5I, and the angle $S\cdots$I-C is 165°. In the alternative binding mode (shown in Fig. 3b), the C_6H_5I molecule is rotated 180° and the iodine atom contacts the carbonyl oxygen atom of Leu84 (distance of 3.4 Å). On the basis of crystallographic refinement, the occupancy of the two binding modes are approximately equal. In the iodopentafluorobenzene-binding mode, the *B*-factors of the iodine atom of C_6H_5I and the S^{δ} atom of Met102 are 22 $Å^2$ and 19 $Å^2$, respectively. In the alternative mode, the *B*-factor of the iodine atom is 27 $Å^2$ and of the oxygen atom of the carbonyl of Leu84 is 20 $Å^2$. With respect to the benzene-binding site (Fig. 1a), the center of the phenyl ring of the C_6H_5I shifts 0.8 Å away from S^{δ} for the iodopentafluorobenzene-binding mode and 1.6 Å towards S^{δ} for the alternative mode (Table 2).

Bromopentafluorobenzene (C_6BrF_5)

The difference electron-density map (Fig. 4), together with crystallographic refinement, suggest

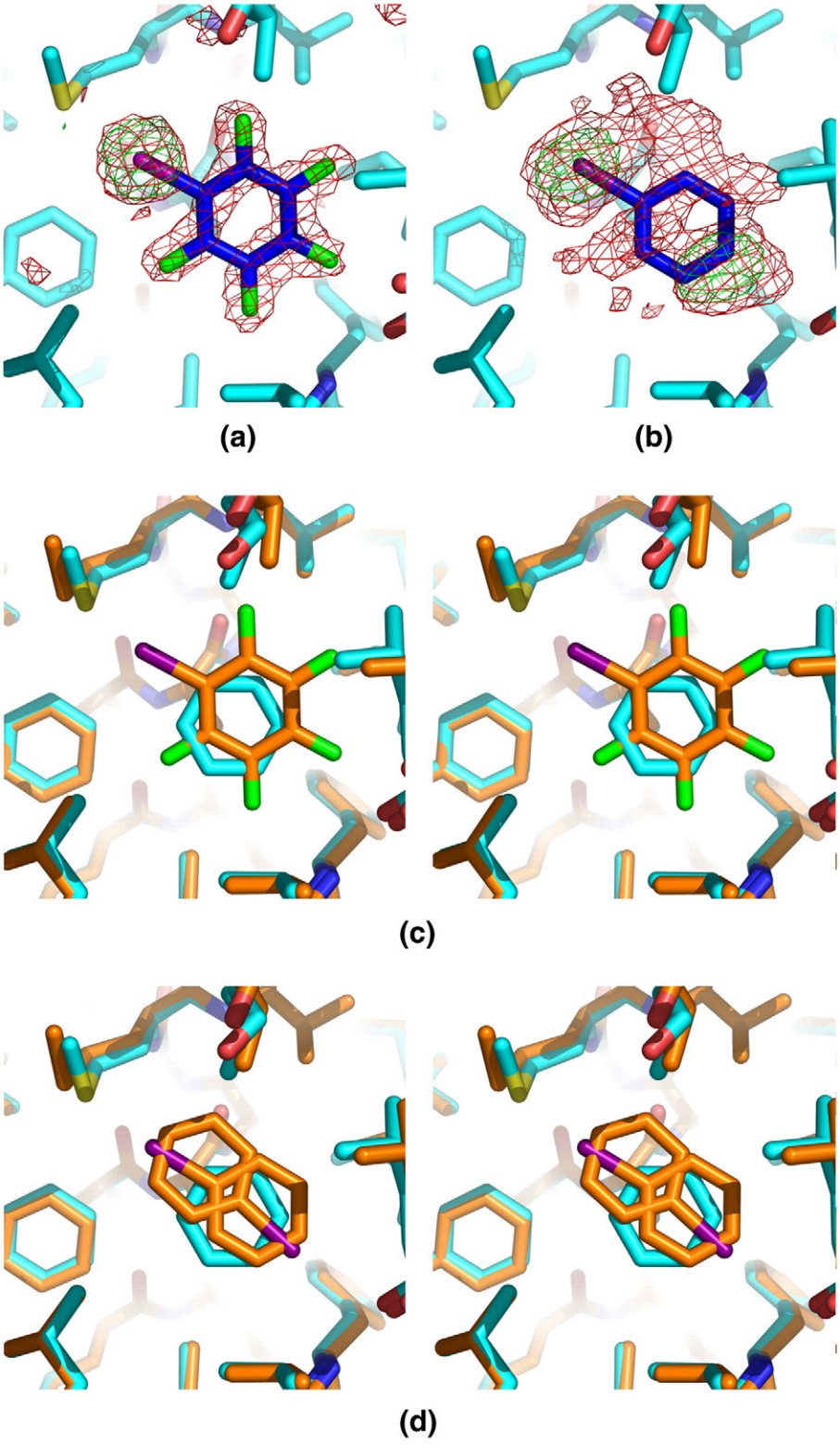

Fig. 3. Binding of C_6F_5I and C_6H_5I. (a) Binding of C_6F_5I. The contours in red are for the ($F_{obs,complex} - F_{obs,L99A}$) isomorphous difference map, calculated as in Fig. 1a and contoured at 3.0σ. The contours in green are for the anomalous difference map, contoured at 3.0σ. The iodine atom is shown in red. (b) Binding of C_6H_5I. The contours in red are from the isomorphous difference map, calculated as in Fig. 1a and contoured at 2.5σ. The contours in green are from the anomalous difference map, contoured at 3.0σ. This map has amplitudes ($F_+ - F_-$), where F_+ and F_- are the Friedel-related amplitudes from the C_6H_5I complex; phases are from the refined structure of L99A. The bound ligand has two binding modes, the first as shown and the second (not shown) with the ligand rotated essentially 180° about an axis coming out of the page. (c) Stereo figure superimposing the complex of C_6F_5I (carbon atoms orange, fluorine green, iodine dark red) on that for C_6H_6 (in blue). The sulfur of Met102, with which the iodine makes a 3.0 Å contact, is at the top left. (d) Stereo figure showing the two modes of binding of C_6H_5I (carbon atoms in orange, iodine in dark red) superimposed on bound C_6H_6 (blue atoms).

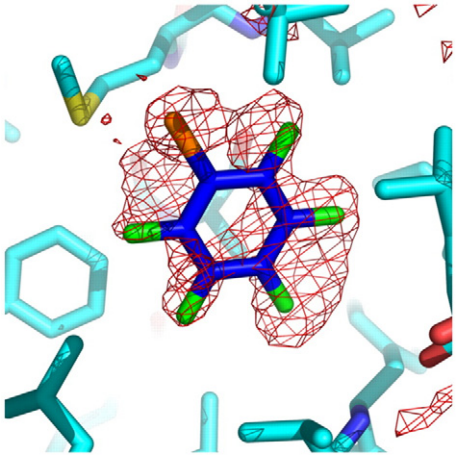

Fig. 4. Isomorphous difference map, calculated as in Fig. 1a showing the binding of C_6BrF_5 to L99A. The bromine atom (orange) is disordered among at least three sites (see the text).

that at least three binding modes can be assigned for C_6BrF_5. In the hexafluorobenzene-binding mode (Fig. 4), the bromine atom is 3.2 Å from S^{δ} of Met102 and the neighboring F···S^{δ} distance is 3.8 Å. In the alternative iodobenzene-binding mode-II (see Fig. 3b), the bromine atom is 3.8 Å from the carbonyl oxygen atom of Leu84, while two fluorine atoms are 3.7 Å and 3.3 Å, respectively, from S^{δ} of Met102. In the third (iodopentafluorobenzene) binding mode (see Fig. 3a), the bromine atom is 3.0 Å from Met102 S^{δ} and the S^{δ}…Br-C angle is 158°.

Binding modes of C_6H_5I and C_6F_5I in the cavity of Se-L99A

One of the most intriguing results from the above studies is the very short I···S contact observed for the C_6F_5I and C_6H_5I complexes. To examine this further, we made the selenomethionine analog of L99A (Se-L99A) and examined its complexes with the same iodine-containing ligands. In Se-L99A all five methionine residues, including Met102, are replaced with selenomethionine. The van der Waals radius of selenium is only ~0.1 Å larger than that of sulfur (Table 3) but its increased electron density and distinct anomalous scattering signal allow its position to be determined more reliably.

The isomorphous difference map and anomalous difference map for C_6F_5I bound to Se-L99A are both shown in Fig. 5a. The mode of binding is very similar to that for the ligand bound to L99A. In particular, the I…Se distance is 3.0 Å and the Se···I-C angle is 169°, both values very close to those observed in the L99A–C_6F_5I complex.

The electron density maps for C_6H_5I (Fig. 5b) show that this ligand also binds to Se-L99A in a manner similar to that for L99A. The I···Se distance in the iodopentafluorobenzene-binding mode of C_6H_5I is 3.4 Å, while the Se···I-C angle is 167°, close to the values observed in the L99A–C_6H_5I complex.

Cavity tolerance and conformational change

The cavity volume in the structure of L99A refined at 100K was determined to be 169 $Å^3$, as compared with 178 $Å^3$ at 273 K.[6] If the side chain of Leu99 in the refined model of WT* at 100K is truncated to alanine, the resultant cavity volume is 167 $Å^3$. The size of the L99A cavity in the refined seleno-Met L99A structure at 100 K was 172 $Å^3$.

Relative to the binding of benzene, all the halogenated benzenes show a significant increase in the cavity volume, ranging from 35 $Å^3$ for C_6H_5I to 64 $Å^3$ for C_6F_5I (Table 2). The van der Waals volume of the halogenated benzenes increases in the order C_6HF_5-C_6F_6<C_6F_5Cl<C_6F_5Br<C_6F_5I (Table 2). However, the observed net increase in cavity volume does not follow this sequence (Table 2). For example, the change of cavity volume in binding C_6HF_5, C_6F_6 and C_6F_5Cl is related approximately to their van der Waals volumes, but the cavity volume has a smaller increase when C_6F_5Br is bound. A change of cavity volume is dependent on ligand size, shape and binding mode.

Binding affinities

To obtain further insights, the affinities of binding of C_6H_6, C_6H_5I and C_6F_5I to L99A and Se-L99A were determined using isothermal titration. As shown in Table 3, C_6H_5I binds most tightly of the three. Also, the affinity of binding of a given ligand to the methionine-containing protein is largely the same as to the selenomethionine-containing protein.

Discussion

Nature of the binding cavity

The L99A cavity is enclosed by amino acid residues with hydrophobic side chains and is almost entirely apolar. It has the general shape of a benzene molecule plus a bulge close to Met102. The part of the protein enclosing the benzene-shaped region is relatively rigid, while the cavity wall in the vicinity of Val111 is more flexible.[3,4] Among the atoms in the wall of the cavity, the sulfur atom of Met102 is the largest and the

Table 3. Affinity of ligand binding measured by isothermal titration calorimetry

Ligand	K_a (M^{-1})	ΔH (kcal/mol)	ΔS (cal/mol/K)
A. *Binding to L99A*			
C_6H_6	17100	−2.5	10.2
C_6H_5I	51800	−4.6	4.9
C_6F_5I	3330	−2.6	6.7
B. *Binding to Se-L99A*			
C_6H_6	11000	−4.4	2.5
C_6H_5I	40900	−5.0	3.1
C_6F_5I	3080	−3.9	1.8

Data were collected at 276 K. K_a is the association constant for ligand binding to protein to yield a protein–ligand complex. ΔH and ΔS are binding enthalpies and entropies for that reaction.

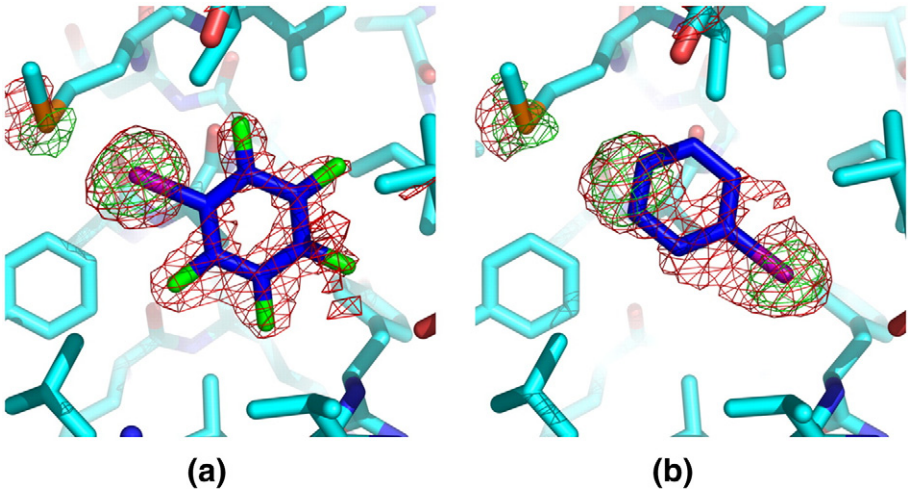

Fig. 5. (a) Maps showing the binding of C_6F_5I to selenium-substituted L99A. The isomorphous difference map has amplitudes ($F_{obs,complex} - F_{obs,L99A}$) and phases from the refined structure of L99A. It is shown in red, and contoured at 3.0σ. C_6F_5I has density for the bound ligand and, due to the replacement of Met102, with the more electron-dense selenomethionine (top left). The map based on the anomalous scattering differences, contoured in green, is contoured at 3.0σ. (b) Electron density maps, as in panel a, for C_6H_5I bound to selenomethionine L99A. The ligand occupies the position shown as well as one rotated 180°.

most polarizable. Main chain atoms are mostly excluded from direct exposure to the cavity, except for the carbonyl oxygen of Leu84 as well as the α-carbon and amide of Tyr88. Both of these residues are opposite the bulge.

Because fluorine is more electron-dense than hydrogen, its position (if ordered) should be seen more clearly in electron density maps. This is exemplified in the map for C_6F_6 (Fig. 1b) relative to that for C_6H_6 (Fig. 1a). In this case, the configuration adopted by the fluorinated ligand is clearly seen. Furthermore, it is obvious that the ligand has a strongly preferred mode of binding and is not freely rotating in the plane of the ring. This "static" mode of binding could occur either because the shape of the cavity is complementary to the ligand, and sterically restricts rotation, or because of chemical complementarity, i.e. favorable interactions between the fluorine atoms on the ligand and specific sites within the cavity. Typically, when a combination of different halogen substituents is present on the benzene ring, and their size does not vary too much (e.g., C_6HF_5 or C_6ClF_5) the ligand is statistically disordered between rotationally distinct modes of binding. If chemical interactions were dominant, one would expect ligands such as these to have a single preferred mode of binding. Because this is not observed, it appears, at least for these ligands, that it is the shape of the cavity and/or the highly apolar chemical environment that predominantly determine the geometry of binding.

F···H-C interactions

As fluorine is the most electronegative atom, the C–F bond is highly polarized and can be considered as a potential hydrogen bond acceptor. Evidence from small-molecule crystal structures suggests that the contribution of F···H-C hydrogen bonding to crystal packing is generally weak. In a study of crystal structures of fluorobenzenes, the intermolecular C···F distance ranged from 3.32 Å to 3.65 Å, suggesting the existence of F···H-C hydrogen bonds, provided that the acidity of the H atom was enhanced and no other competing acceptors were present.[7] The T4 L99A cavity does offer such an environment.

For different ligands, the number of F···H-C contacts varied from about eight to more than 20, in a highly model-dependent manner. The shortest F···C distances are around 3.1 Å, while others range up to 3.5 Å. The carbon atoms engaged in the F···H-C contacts are mostly contributed by the side chains of the residues in the cavity, except for a few C^{α} atoms. The α-carbon of Tyr88, in particular, has close contacts to a fluorine atom of most ligands. This tyrosine residue is located

Table 4. Radii and polarizability data for representative atoms and bonds

Atom X	van der Waals radius[a] (Å)	Atomic polarizability[b] (Å^3)	C–X bond polarizability[c] (Å^3)
H	1.1	0.67	0.664
F	1.35	0.56	0.575
Cl	1.80	2.18	2.580
Br	1.95	3.05	3.722
I	2.15	5.35	5.791
C	1.70	1.76	0.514
S	1.85	2.90	N/A
Se	2.00	3.77	N/A
O	1.40	0.80	0.161 (C–O)
O	1.40	0.80	1.383 (C═O)
N	1.5	1.10	0.611

[a] Ref. 36.
[b] Data from Ref. 37.
[c] Ref. 34 and references therein.

opposite to Met102 in a highly ordered part in the cavity that defines the limit to which ligands can deform the protein structure. Whenever the C^{α} atom of Tyr88 contacts a fluorine atom, the NH group contacts the same fluorine (distance ~3.1–3.5 Å). The F···H-N contact appears to be around 0.1–0.2 Å shorter than the F···H-C contact in all cases, perhaps indicative of a stronger interaction, but also possibly due to the smaller van der Waals radius (Table 4).

All the side chains within the cavity are apolar and hydrophobic, implying low acidity of their hydrogen atoms. Although the many F···C contacts fall into the range of potential F···H-C hydrogen bonds, the variability in ligand-binding modes suggests that hydrogen bonding is not critical. Rather, van der Waals interactions are presumably of primary importance in defining the binding configuration.

The expected length of a C-F bond is 1.34 Å, and the van der Waals radius of fluorine itself is 1.35 Å (Table 4). For hydrogen, the corresponding values are both 1.1 Å. Therefore, a hydrogen-to-fluorine replacement on a phenyl ring will produce an approximately 0.5 Å difference in the van der Waals contact distance relative to the carbon atom. These changes are reflected in the modes of binding observed for C_6H_6 and C_6F_6 (Fig. 1d). Relative to benzene, the centroid of the fluorinated ligand moves 1.0 Å towards Met102 and the ring has a largely out-of-plane rotation of 21° (Table 2). In the process, the volume of the cavity increases by 52 $Å^3$, mostly through movements of the relatively flexible F helix.[4]

Iodine-sulfur and iodine–selenium interaction

Perhaps the most striking findings of this study are the remarkably close iodine–sulfur and iodine–selenium contacts. This was first seen most clearly in the complex of C_6F_5I with L99A (Fig. 3a and 3c). X-ray refinement indicated that the distance from the iodine on the ligand to the sulfur in the side chain of Met102 was 3.0 Å, which is 1.0 Å less than the sum of the respective van der Waals radii (Table 3). There could be some concern that the high electron density of the iodine atom might cause some uncertainty in the positions of the neighboring, less electron-dense atoms, including the sulfur.

To provide further evidence, we collected data for ligand complexes with the L99A protein in which all methionines, including Met102, were substituted with selenomethionine. The selenium atom is more electron dense than sulfur, and so can be located with greater precision. In addition, the strong anomalous scattering of both selenium and iodine provides a powerful check on the positions of these two atoms.

The results for the selenium-substituted L99A lysozyme fully support those for the protein with methionine. In particular, in the complex with C_6F_5I (Fig. 5a) the distance between the iodine atom and the selenium atom in the SeMet102 side chain is 3.0 Å, 1.15 Å less than the sum of the van der Waals radii. C_6H_5I binds to selenium-substituted L99A in the same disordered fashion as to L99A itself (Fig. 5b). The closer iodine is 3.4 Å from the selenium.

Halogen-bond formation

Interactions of the form C-X···B, where X = F, Cl, Br and I and B = O, N or S have been fairly well characterized, and are referred to as halogen bonds.[8-11] Particularly when X is a substituent on an aromatic ring, polarization along the C-X bond results in a positive potential on the halogen.[9] The energy of a halogen bond increases with the polarizability of the halogen, being greatest for iodine. Fluorines rarely participate in such bonds.[12] Halogen bonding is characterized by charge-transfer or donor–acceptor interaction and has recently been viewed as so-called σ-hole bonding.[13-15] There are a number of similarities between halogen bonds and hydrogen bonds and they can be competitive in some situations.

The present results suggest clearly that the iodine in C_6F_5I forms a halogen bond with the sulfur of Met102. The iodine-sulfur distance is 3.0 Å which is about 0.6 Å longer than the sum of their covalent radii and 1.0 Å less than the sum of their van der Waals radii (Table 3). The angle S^{δ}···I-C is 166° which is reasonably close to the canonical linear geometry.[8] The C^{γ}-S^{δ}···I and C^{ε}-S^{δ}···I angles are 104° and 111°, respectively. All three angles fall into the regions that are most favored for halogen bonds.[9] In addition, two angles, ϕ and φ, defining the directionality of the halogen bond,[8] are determined to be 3.5° and 107°, respectively. Again, these values are in a good agreement with those observed from small molecular crystallography.[8]

Likewise, the evidence is also clear that there is a halogen bond between the iodine in C_6F_5I and the selenium of SeMet102. The geometry defining this halogen bond is very similar to that between C_6F_5I and Met102. Du Mont *et al.*[16] (see also Ref. 17) have pointed out that selenium–iodine contacts range from van der Waals contacts through secondary soft–soft interactions (2.65–2.95 Å) to covalent bonds (~2.5 Å). The distance of 3.0 Å that we observe is at the boundary between a "secondary bond" and a Se–I interaction approaching a bond order of 0.5. It might be noted also that a Se–I interaction in the enzyme–substrate complex is thought to facilitate the *in vivo* deiodination of the prohormone thyroxine to its biologically active form.[18,19]

The evidence for halogen bond formation in the compounds other than C_6F_5I is not especially compelling. For example, C_6H_5I displays two different modes of binding, one with the iodine atom close to S^{δ} of Met102 (3.3 Å), and the other with the iodine 8.3 Å away (Fig. 3b). In the second mode of binding, the iodine is 3.4 Å from the carbonyl oxygen of Leu84 (~0.2 Å less than the sum of the van der Waals radii). In addition, the angles O···I–C and C═O···I are 154° and 120°, which would be consistent with a possible O···X type halogen bond.[9] On the other hand, C_6H_5I is smaller than C_6F_5I (Table 2) and is presumably free to adopt the same (single) mode of binding. That it does not do so suggests that the apparently favorable interaction between the iodine and sulfur occurs for C_6F_5I but not to the same degree for C_6H_5I. For C_6F_5I, the strong electronegativity of the five fluorine atoms will influence the behavior of atom X. The hydrogen

atom of C_6HF_5, for example, shows much higher acidity ($pK_a = 24.2$) than that of C_6H_6 ($pK_a = 43$).[20] Thus, the iodine atom in C_6F_5I will tend to make a stronger halogen bond than is the case for C_6H_5I. This is consistent with the observed differences in binding behavior.

Relative to benzene, the binding affinity of C_6H_5I is significantly increased for both L99A and Se-L99A. Taken together with the observation that the cavity needs to expand more to accommodate the latter ligand (Table 2), it clearly indicates that the I···S and I···Se interactions are favorable. Notwithstanding the short I···S and I···Se approaches, and the single conformation seen for C_6F_5I bound to L99A and Se-L99A, this ligand binds less tightly than C_6H_5I (Table 3), which has two distinct binding modes (Fig. 3b). This difference might be rationalized in terms of simple size considerations; i.e. the volume of C_6F_5I is about 35% larger than that of C_6H_5I (Table 2), and the cavity is required to expand by an equivalent amount to accommodate the former compound, which could explain its weaker binding.

For X = H, Cl and Br, the binding of C_6F_5X shows disorder with atom X occupying two or more alternative sites (Figs. 1c, 2a, and 4). This tends to suggest that neither the chloride or the bromide atom has a strong tendency to form a halogen bond with the sulfur of Met102 in the cavity wall. It is consistent also with the expectation that a halogen bond for chlorine and bromine is weaker than that for iodine.[8-10,13,14]

Recently, Voth *et al.*[21] have used competing crystal structures of DNA hybrids to suggest that the strength of a bromine-oxygen halogen bond is ~2-5 kcal/mol stronger than the analogous hydrogen bond (or more when the halogen is buried). We cannot make exactly the same comparison but we can, for example, compare the binding of C_6H_5I, which does make a halogen bond, with C_6H_6, which does not. Using the relationship $\Delta G° = -RT \ln K_a$, the difference in free energies of binding of these ligands to L99A or to Se-L99A (Table 3) is 0.6–0.7 kcal/mol. A better comparison is between C_6H_5I and toluene, since these two ligands have essentially identical size and shape. The association constant of toluene to L99A [22] at 29 °C is 9800 M^{-1}, which is 0.5 kcal/mol weaker than that at C_6H_5I (at 3 °C). These values of 0.5–0.7 kcal/mol are substantially less than the values of 2–5 kcal/mol suggested by Voth *et al.*,[21] even though halogen bonds involving iodine are thought to be stronger than those with bromine. (As noted above, in the present experiments neither the chlorine nor bromine-containing ligands show evidence for significant halogen bonding.) It should also be noted that the values of 0.5–0.7 kcal/mol are for a halogen bond relative to a van der Waals contact, rather than relative to a hydrogen bond. Given that a hydrogen bond is typically 1–2 kcal/mol stronger than a van der Waals contact, the present results suggest that the halogen bond in the T4 lysozyme apolar cavity might actually be weaker than a hydrogen bond. These apparent differences with the study of Voth *et al.*[21] still need to be resolved.

Materials and Methods

Protein purification, crystallization and protein-ligand complexes

The T4 lysozyme mutant L99A was constructed in the cysteine-free pseudo wild type protein following previously reported methods.[23] Crystals isomorphous with wild type in space group $P3_221$ were obtained by the hanging-drop method at 4 °C with the standard precipitant solution (2.0–2.4 M sodium/potassium phosphate, pH 6.7–7.1, 50 mM β-mercaptoethanol (β-ME) and 50 mM hydroxyethyldisulfide) as described.[23] The selenomethionine-substituted protein (Se-L99A) was produced by the feedback inhibition method of van Duyne *et al.*[24] and purified as above in the presence of 10 mM L-methionine. Crystals of Se-L99A were obtained with the same protocol.

The protein–ligand complexes were obtained by vapor-diffusion, mimicking the hanging-drop method. The ligands were kept well sealed as far as possible and were handled in a hood with good ventilation. The protein crystal soaked in 50 μl of precipitant was arranged as a hanging-drop, and the pre-cooled (4 °C) ligand was placed in an aluminum or glass boat and floated on the reservoir solution. This avoided direct contact between the crystals and the bulk ligands. For the liquid [C_6F_5X (X = H, F, Cl, Br, or I), C_6H_6 and C_6H_5I] and solid ($C_6F_3Cl_3$, $C_6F_4Br_2$ and $C_6F_4I_2$) compounds, overnight at 4 °C and three days at 4 °C, respectively, were allowed for the ligand to diffuse and reach equilibrium. Before data collection, crystals were soaked in a cryoprotectant-containing solution (precipitant solution plus 18% glycerol) for 5 s and flash-frozen in liquid nitrogen and then moved into the cryogenic nitrogen stream (100 K). To avoid the loss of ligands during cryosoaking, the cryoprotectant-containing solution was equilibrated overnight before use in the same manner as for producing the complexes with the various ligands.

X-ray diffraction and structure refinement

X-ray diffraction experiments were performed with a Rigaku rotating anode X-ray generator (λ = 1.54 Å). Data were collected on an Raxis IV imaging plate detector. Ligand-free data were collected to 1.6 Å resolution, whereas those for complexes were collected to 1.8 Å resolution. Data were processed with Denzo/Scalepack.[25] Structures were refined with Refmac[26] using protocols established for T4 lysozyme mutants[27,28] with 5% of reflections randomly selected and used for free *R*-factor calculation. The occupancy of ligands with multiple conformations was refined with CNS.[29] Modeling was performed with O.[30] The previously reported ligand-free L99A structure (PDB entry 1L90) at room temperature was used as the initial model, and redetermined at 100 K. Rigid-body refinement was executed first, followed by simulated annealing and energy minimization. This structure was used as the reference in calculation of ($F_{obs,complex} - F_{obs,L99A}$) difference maps. These maps were used to obtain starting models for the ligands. The expected structural parameters for the ligands were obtained from the Cambridge Structural Database and were used as restraints in refinement. Water molecules, phosphate groups and β-ME were added based on $F_{obs} - F_{calc}$ difference maps with a 3σ threshold. Further positional and restrained *B*-factor refinements were performed for all atoms, and several cycles of occupancy refinement were applied when appropriate for those complexes with multiple ligand-binding modes.

Calculation of cavity volume and area

Cavity volumes and areas were calculated with MSP version 3.7[31] as described.[6] For consistency, the default atomic radii and a probe radius of 1.4 Å were applied.

Ligand binding

Binding constants and enthalpies of binding were estimated using a VP-ITC isothermal titration calorimeter (MicroCal, Northampton, MA). Operating in multiple injection mode, a 1–2 mM solution of Se-L99A or L99A was injected stepwise into buffer to determine a baseline injection heat profile, and then into buffer containing ligands dissolved overnight to about 0.1 mM with stirring on ice. A 1 μl, 28 × 10 μl injection scheme with 180 s between injections was used. Baseline heats were subtracted from reaction heats and the resulting profile fit to a single-site binding model using Origin version 7E (OriginLab Corp., Northampton, MA). Titrations were done at 3 °C with stirring at 307 rpm. The buffer was 50 mM sodium acetate (50 mM $H_{0.14}Na_{0.86}OAc$), pH 5.45.[3]

Proteins were dialyzed into outgassed buffer. Concentrations were determined with an HP-8453 spectrophotometer (Agilent Technologies, Santa Clara, CA) and the molar absorptivity at 280 nm of 24,170 $L \cdot mol^{-1} \cdot cm^{-1}$) for T4 lysozyme determined by Elwell & Schellman.[32]

In modeling the binding reaction, ligand concentrations were attenuated to give unit binding stoichiometries. Reductions of 10–15% in nominal concentrations were usual and most likely due to the high degree of volatility of these compounds.

Protein Data Bank accession codes

The coordinates and structure amplitudes have been deposited in the Protein Data Bank. Accession codes (Table 1) are 3DMV, 3DMX, 3DMZ, 3DN0, 3DN1, 3DN2, 3DN3, 3DN4, 3DN6, 3DN8, and 3DNA.

Acknowledgements

We thank Andy Fields for making the L99A protein. This work was supported, in part, by NIH grant GM21967 to B.W.M.

References

1. Eriksson, A. E., Baase, W. A., Zhang, X.-J., Heinz, D. W., Blaber, M., Baldwin, E. P. & Matthews, B. W. (1992). Response of a protein structure to cavity-creating mutations and its relation to the hydrophobic effect. *Science*, **255**, 178–183.
2. Eriksson, A. E., Baase, W. A., Wozniak, J. A. & Matthews, B. W. (1992). A cavity-containing mutant of T4 lysozyme is stabilized by buried benzene. *Nature*, **355**, 371–373.
3. Morton, A., Baase, W. A. & Matthews, B. W. (1995). Energetic origins of specificity of ligand binding in an interior nonpolar cavity of T4 lysozyme. *Biochemistry*, **34**, 8564–8575.
4. Morton, A. & Matthews, B. W. (1995). Specificity of ligand binding in a buried nonpolar cavity of T4 lysozyme: Linkage of dynamics and structural plasticity. *Biochemistry*, **34**, 8576–8588.
5. Wei, B. Q., Baase, W. A., Weaver, L. H., Matthews, B. W. & Shoichet, B. K. (2002). A model binding site for testing scoring functions in molecular docking. *J. Mol. Biol.* **322**, 339–355.
6. Quillin, M. L., Breyer, W. A., Griswold, I. J. & Matthews, B. W. (2000). Size *versus* polarizability in protein-ligand interactions: binding of noble gases within engineered cavities in phage T4 lysozyme. *J. Mol. Biol.* **302**, 955–977.
7. Thalladi, V. R., Weiss, H. C., Blaser, D., Boese, R., Nangia, A. & Desiraju, G. R. (1998). C-H···F interactions in the crystal structures of some fluorobenzenes. *J. Am. Chem. Soc.* **120**, 8702–8710.
8. Ouvrard, C., Le Questel, J.-Y., Berthelot, M. & Laurence, C. (2003). Halogen-bond geometry: a crystallographic data-base investigation of dihalogen complexes. *Acta Cryst.* **B59**, 512–526.
9. Auffinger, P., Hays, F. A., Westhof, E. & Ho, P. S. (2004). Halogen bonds in biological molecules. *Proc. Natl Acad. Sci. USA*, **101**, 16789–16794.
10. Metrangolo, P., Neukirch, H., Pilati, T. & Resnati, G. (2005). Halogen bonding based recognition processes: a world parallel to hydrogen bonding. *Acc. Chem. Res.* **38**, 386–395.
11. Metrangolo, P. & Resnati, G. (2008). Halogen versus hydrogen. *Science*, **321**, 918–919.
12. Politzer, P., Murray, J. S. & Concha, M. C. (2007). Halogen bonding and the design of new materials: organic bromides, chlorides and perhaps even fluorides as donors. *J. Mol. Model.* **13**, 643–650.
13. Clark, T., Hennemann, M., Murray, J. S. & Politzer, P. (2007). Halogen bonding: the sigma-hole. *J. Mol. Model.* **13**, 291–296.
14. Politzer, P., Murray, J. S. & Lane, P. (2007). σ-Hole bonding and hydrogen bonding: competitive interactions. *Int. J. Quantum Chem.* **107**, 3046–3052.
15. Murray, J. S., Lane, P. & Politzer, P. (2008). Simultaneous s-hole and hydrogen bonding by sulfur- and selenium-containing heterocycles. *Int. J. Quantum Chem.* **108**, 2770–2781.
16. du Mont, W. -W., Martens-von Salzen, A., Ruthe, F., Seppala, E., Mugesh, G., Devillanova, F. A. *et al.* (2001). Tuning selenium-iodine contacts: from secondary soft-soft interactions to covalent bonds. *J. Organometallic Chem.* **623**, 14–28.
17. Klapotke, T. & Passmore, J. (1989). Sulfur and selenium iodine compounds: From nonexistence to significance. *Acc. Chem. Res.* **22**, 234–240.
18. Berry, M. J., Banu, L. & Larsen, P. R. (1991). Type I iodothyronine deiodinase is a selenocysteine-containing enzyme. *Nature*, **349**, 438–440.
19. Berry, M. J., Kieffer, J. D., Harney, J. W. & Larsen, P. R. (1991). Selenocysteine confers the biochemical properties characteristic of the Type I iodothyronine deiodinase. *J. Biol. Chem.* **266**, 14155–14158.
20. Brown, J. M. & Chaloner, P. A. (1977). Strong amide-halothane hydrogen-bonding observed by nuclear magnetic resonance. *Can. J. Chem.* **55**, 3380–3383.
21. Voth, A. R., Hays, F. A. & Ho, P. S. (2007). Directing macromolecular conformation through halogen bonds. *Proc. Natl Acad. Sci. USA*, **104**, 6188–6193.
22. Morton, A., Baase, W. A. & Matthews, B. W. (1995). Energetic origins of specificity of ligand binding in an interior nonpolar cavity of T4 lysozyme. *Biochemistry*, **34**, 8564–8575.
23. Eriksson, A. E., Baase, W. A. & Matthews, B. W. (1993). Similar hydrophobic replacements of Leu 99 and Phe

153 within the core of T4 lysozyme have different structural and thermodynamic consequences. *J. Mol. Biol.* **229**, 747–769.
24. Van Duyne, G. D., Standaert, R. F., Karplus, P. A., Schreiber, S. L. & Clardy, J. (1993). Atomic structures of the human immunophilin FKBP-12 complexes with FK506 and rapamycin. *J. Mol. Biol.* **229**, 105–124.
25. Otwinowski, Z. & Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326.
26. CCP4: Collaborative Computational Project Nr4. (1994). The CCP4 suite: Programs for protein crystallography. *Acta Crystallogr. D*, **50**, 760–763.
27. Baldwin, E., Baase, W. A., Zhang, X.-J., Feher, V. & Matthews, B. W. (1998). Generation of ligand binding sites in T4 lysozyme by deficiency-creating substitutions. *J. Mol. Biol.* **277**, 467–485.
28. Dao-pin, S., Alber, T., Baase, W. A., Wozniak, J. A. & Matthews, B. W. (1991). Structural and thermodynamic analysis of the packing of two α-helices in bacteriophage T4 lysozyme. *J. Mol. Biol.* **221**, 647–667.
29. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W. *et al.* (1998). Crystallography and NMR system (CNS): A new software system for macromolecular structure determination. *Acta Crystallogr. D*, **54**, 905–921.
30. Jones, T. A. (1982). FRODO: a graphics fitting program for macromolecules. In *Crystallographic Computing* (Sayre, D., ed.), pp. 303–317, Oxford University Press, Oxford, UK.
31. Connolly, M. L. (1993). The molecular surface package. *J. Mol. Graph.* **11**, 139–141.
32. Elwell, M. & Schellman, J. (1975). Phage T4 lysozyme physical properties and reversible unfolding. *Biochim. Biophys. Acta*, **386**, 309–323.
33. Miller, K. J. & Savchik, J. A. (1979). A new empirical method to calculate average molecular polarizabilities. *J. Mol. Biol.* **101**, 7206–7213.
34. Miller, K. J. (1990). Additivity methods in molecular polarizability. *J. Am. Chem. Soc.* **112**, 8533–8542.
35. Zhao, Y. H., Abraham, M. H. & Zissimos, A. M. (2003). Fast calculation of van der Waals volume as a sum of atomic and bond contributions and its application to drug compounds. *J. Org. Chem.* **68**, 7368–7373.
36. Pauling, L. (1970). *General Chemistry*, 3rd edit. W.H. Freeman & Co., San Francisco.
37. Lide, D. R. (1996). *CRC Handbook of Chemistry and Physics*, 77th edit. CRC Press, Boca Raton.