

Structures de la 5'-Bromo-5'-deoxythymidine

PAR M. HUBER*

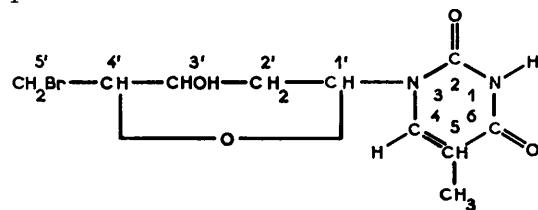
Crystallographic Laboratory, Cavendish Laboratory, Cambridge, England

(Reçu le 20 juillet 1956)

The structure of 5'-bromo-5'-deoxythymidine has been determined for each of two crystalline modifications, by two-dimensional projections and generalized projections. Accurate measurement of bond lengths has not been attempted. The ribose and thymine residues are in the standard configuration. From the system of hydrogen bonds, some conclusions are drawn about the tautomeric form of the molecule.

Introduction

L'étude de la structure de ce composé, de formule proposée:



a été entreprise pour confirmer les connaissances obtenues par voie chimique, et en particulier la configuration du sucre, et les positions relatives des deux parties de la molécule.

Dans les composés du même type, le sucre est sous forme cyclique, la liaison C–N entre les deux cycles est du type β et les plans moyens des deux cycles sont à angle droit.

Dans le cas présent, il restait, entre autres choses, à préciser les particularités du cycle déoxyribofuranose.

Étude expérimentale

Les cristaux utilisés ont été aimablement fournis par le Dr Michelson. Ils ont été obtenus à partir d'une solution aqueuse et se présentaient sous forme d'aiguilles de 0,5 à 1 mm. de long et de 0,05 à 0,1 mm. transversalement.

Deux formes ont été décelées dans le produit cristallisé. L'une *A* très plastique, et sans faces extérieures définies, est monoclinique, l'autre *B*, présentant des facettes bien formées, tant latérales que terminales, est orthorhombique.

Les diagrammes de rayons X, pour lesquels les temps de pose s'étagaient entre 10 minutes et 30 heures, ont été enregistrés photographiquement dans une chambre de Weissenberg, avec la radiation $Cu K\alpha$ filtrée par le nickel.

Les dimensions et les caractéristiques des mailles sont:

* Maintenant au Laboratoire de Chimie Minérale B, 11 rue Pierre Curie, Paris 5, France.

Forme *A*.Monoclinique, groupe spatial $B_2-C_2^3$
 $a = 25,50; b = 5,50; c = 8,95 \text{ \AA}; \beta = 108^\circ \pm 1^\circ.$
 (Longueurs définies à $\pm 0,05 \text{ \AA}$.)
Volume 1 $200 \pm 20 \text{ \AA}^3$; 4 molécules par maille.**Forme *B*.**Orthorhombique, groupe spatial $P2_12_12_1-D_4^2$
 $a = 23,10; b = 9,55; c = 5,42 \text{ \AA}.$
 (Longueurs définies à $\pm 0,03 \text{ \AA}$.)
Volume 1 $200 \pm 12 \text{ \AA}^3$; 4 molécules par maille.

L'axe court dans les deux cas coïncide avec l'axe de l'aiguille.

La forme *B* présentant un centre de symétrie pour les projections (001) et (010) et dont les cristaux se clivent facilement se prête mieux que la forme *A*, non clivable, et ne possédant de centre de symétrie que pour la projection (010), à une détermination de structure par séries de Fourier; notre étude a porté essentiellement sur cette seconde forme.

Détermination des structures

Les mesures d'intensité par comparaison visuelle avec un étalon de noircissement ont porté sur les zones suivantes:

A: $h0l, h1l, h2l$; B: $hk0, h0l$.

Elles ont fourni les intensités pour quelques 70 % des spectres théoriquement observables. Les intensités ont été corrigées des facteurs de Lorentz et de polarisation, et les différentes zones mises à des échelles comparables, soit en utilisant les axes communs, soit pour les fonctions de Patterson, en utilisant l'altitude du pic origine.

Les fonctions de Patterson ont fourni les coordonnées du brome pour les trois projections étudiées, et l'étude par série de Fourier a comporté les stades suivants:

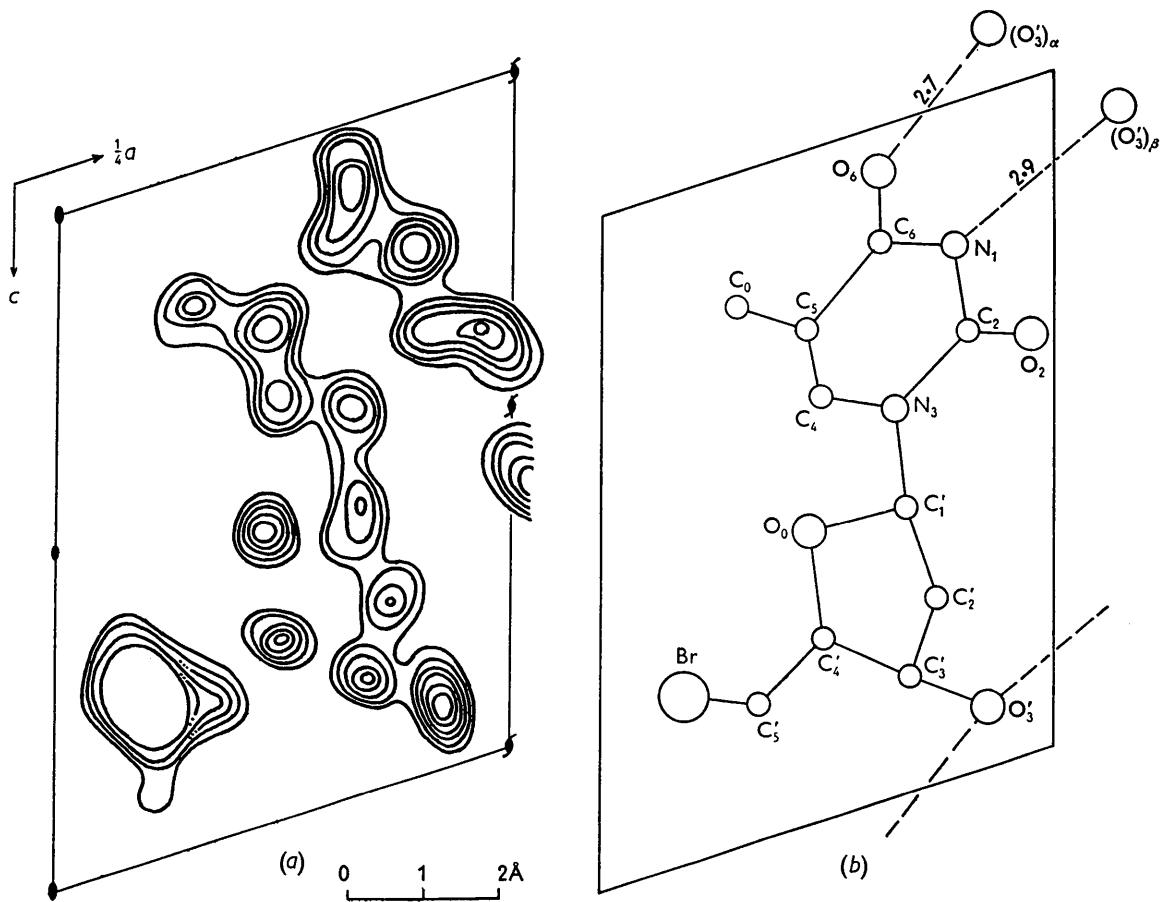


Fig. 1. Forme A. (a) Densité électronique projetée sur (010). (b) Longueurs et angles des liaisons (les liaisons hydrogène en trait interrompu).

Forme A : 3 synthèses $F_o(h0l)$, 3 synthèses $F_o - F_c(h0l)$, synthèses généralisées $h1l$ et $h2l$ de Patterson.

Forme B : 3 synthèses $F_o(hk0)$, 3 synthèses $F_o(h0l)$, 3 synthèses $F_o - F_c(h0l)$.

Seule la forme B a été étudiée avec quelque précision du point de vue des coordonnées spatiales.

Les facteurs de structure F_c ont été calculés à l'aide des valeurs f_0 des facteurs atomiques de structure prises dans les *Tables Internationales* (1935), puis modifiées pour avoir des valeurs comparables aux F_o par un facteur de température.

Les contours d'égal densité électronique ont été tracés à une échelle arbitraire, et ceux-ci sont omis dans l'atome de brome (Figs. 1(a), 2(a) et 2(b)).

Dans les projections A(010) (Fig. 1(a)) et B(001) (Fig. 2(a)) suivant l'axe court, les atomes sont résolus à deux exceptions près, et à elles seules ces projections suffisent à prouver l'identité fondamentale des deux formes du point de vue des configurations et des formules, ainsi qu'à lever les indéterminations majeures en raison desquelles ce travail était entrepris.

La projection B(010) (Fig. 2(b)) par contre ne pré-

sente que deux atomes complètement résolus et n'a pas permis d'obtenir les coordonnées z avec le même ordre de précision que les coordonnées x et y .

Pour la forme A, les fonctions généralisées de Patterson (Dyer, 1951; Cochran & Dyer, 1952) ont permis de déterminer une valeur approximative des coordonnées y pour neuf des atomes de la molécule (coordonnées prises par rapport à $y_{\text{Br}} = 0,000$ par hypothèse), ce qui a permis, grossièrement, en utilisant le modèle fourni par B, de placer la molécule dans la maille et de localiser les atomes restants. Nous n'avons pas jugé utile de poursuivre cette étude par synthèse généralisée de Fourier.

Pour la forme B, l'écart moyen calculé est, pour les atomes C, N, O $\sigma(x) = \sigma(y) = 0,05$ Å et l'erreur sur les z calculés au plus proche $\frac{1}{60}$ ème près est de l'ordre de $\pm 0,04$ Å, et par conséquent l'écart moyen sur la longueur des liaisons est de l'ordre de 0,07–0,10 Å, l'erreur maximale restant inférieure à 0,2 Å. L'erreur correspondante sur les angles de liaison peut atteindre 6–7° (Figs. 1(b) et 2(c)).

Les longueurs de liaison sont en accord avec celles généralement admises, et comme l'objet de cette étude

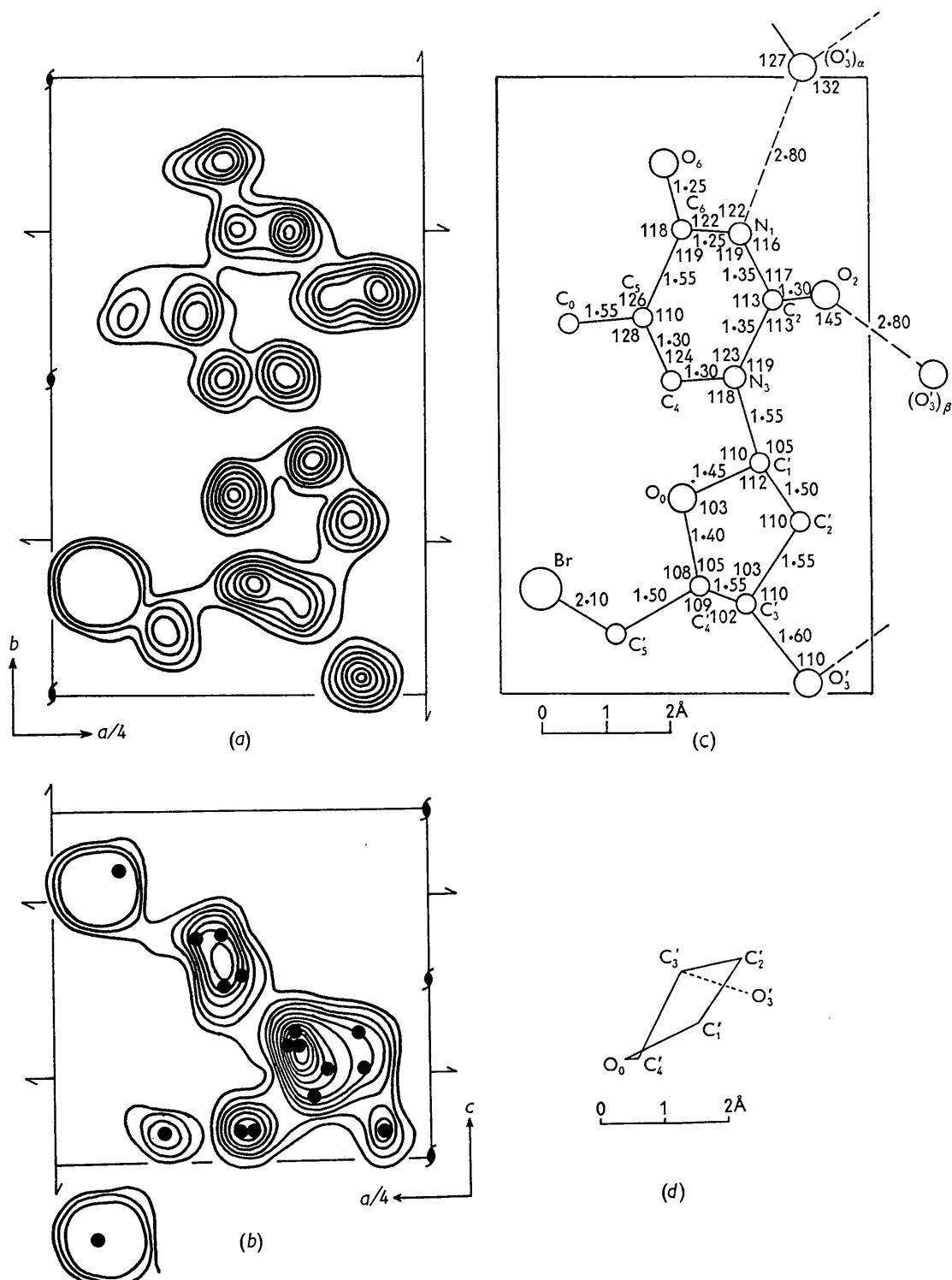


Fig. 2. Forme B. (a) Densité électronique projetée sur (001). (b) Densité électronique projetée sur (010). (c) Longueurs et angles des liaisons. (d) Le cycle pentagonal vu suivant l'axe b .

était de déterminer la stéréochimie de la molécule, nous ne nous sommes pas attachés à réduire l'erreur sur les longueurs de liaison indiquées plus haut.

Pour la forme *A*, nous n'avons pas précisé les longueurs de liaisons, puisque la structure trouvée pour *B* se trouve être en parfait accord avec la projec-

Tableau 1. Coordonnées atomiques fractionnaires

	Forme A				Forme B		
	x	z	y _I	y _{II}	x	y	z
C _{1'}	0,3367	0,0750	—	0,89	0,1732	0,3750	0,183
C _{2'}	0,3750	0,2250	—	0,78	0,2017	0,2817	0,367
C _{3'}	0,3383	0,3250	0,65-0,70	0,69	0,1639	0,1465	0,333
C _{4'}	0,2416	0,2333	—	0,51	0,1333	0,1750	0,083
C _{5'}	0,1750	0,3000	0,30-0,35	0,32	0,0766	0,0991	0,083
C ₂	0,5950	0,1533	0,00-0,10	0,09	0,1843	0,6390	0,266
C ₄	0,7583	0,1166	—	0,36	0,1150	0,5083	0,500
C ₅	0,7750	0,2500	0,50-0,60	0,58	0,0967	0,6135	0,633
C ₆	0,6916	0,3250	—	0,58	0,1227	0,7523	0,533
C ₉	0,8500	0,3000	—	0,76	0,0467	0,6000	0,833
N ₁	0,6116	0,2916	—	0,36	0,1610	0,7467	0,367
N ₃	0,6750	0,0666	0,05-0,10	0,09	0,1578	0,5141	0,333
O ₆	0,2333	0,0700	0,65-0,75	0,64	0,1217	0,3163	0,083
O ₃	0,4283	0,4033	0,55-0,60	0,58	0,2050	0,0173	0,267
O ₂	0,5250	0,1166	0,8-0,9	0,89	0,2197	0,6470	0,083
O ₅	0,6916	0,4500	0,65-0,75	0,76	0,1118	0,8600	0,650
Br	0,9333	0,2516	—	0,000	0,0283	0,1700	0,7833

y_I: valeur déduite des fonctions de Patterson généralisées.*y_{II}*: valeur de *y* présentant le meilleur accord avec les dimensions moléculaires de la forme B

Tableau 2. Forme B: longeur des liaisons et distances interatomiques

Br ₂ -C _{5'}	2,10 Å	C _{3'} -C _{4'}	1,55 Å	C ₅ -C ₆	1,55 Å
C _{5'} -C _{4'}	1,50	C _{3'} -O _{3'}	1,60	C ₆ -N ₁	1,25
C _{4'} -O ₉	1,40	C _{1'} -N ₃	1,55	C ₆ -O ₆	1,25
O ₉ -C ₁	1,45	N ₃ -C ₄	1,30	N ₁ -C ₂	1,35
C ₁ -C ₂	1,50	C ₄ -C ₅	1,30	C ₂ -N ₃	1,35
C ₂ -C ₃	1,55	C ₅ -C ₀	1,55	C ₂ -O ₂	1,30
N ₁ -(O _{3'}) _α	2,80 Å	O ₂ -(O _{3'}) _β	2,80 Å		
O ₆ -(O _{3'}) _α	3,35	O ₃ -(O _{3'}) _γ	3,40		

tion A(010) et se place très bien spatialement dans la maille A, compte tenu des coordonnées *y* qui ont été déterminées. Nous avons pourtant calculé les deux distances

$$O_6-(O_3')_\alpha = 2,7 \pm 0,1 \text{ Å}, N_1-(O_3')_\beta = 2,9 \pm 0,1 \text{ Å}$$

entre atomes de molécules différentes, qui sont intéressantes du point de vue structural.

Pour les deux formes les facteurs $R = \sum |F_o - F_c| / \sum |F_o|$ restent comparativement élevés:

$$R_A(h0l) = 0,27, R_B(hk0) = 0,20, R_B(h0l) = 0,21 \\ (\text{calculés en incluant tous les zéros observés});$$

$$R_A(h0l) = 0,24, R_B(hk0) = 0,15, R_B(h0l) = 0,17 \\ (\text{calculés en excluant les zéros}).$$

Ceci doit être attribué à la mesure assez imprécise des F_o , (60% seulement des F_o pour la forme A ont pu être mesurés), et aussi au fait que le facteur atomique de structure du brome par rapport aux atomes légers n'a pas été ajusté. (Un essai en ce sens, par la méthode des moindres carrés, et portant sur le domaine $0 < \sin \theta < 0,60$ pour la forme A a permis de réduire R de 0,22 à 0,17 pour cette zone.) De plus, pour la forme A la forte anisotropie de vibration du brome est responsable d'une cause supplémentaire de désaccord dont on n'a pas tenu compte.

Un écart important pour le spectre (410)_B entre F_o et F_c est probablement dû à une forte extinction de cette réflexion, puisque la valeur absolue de F_c pour cette réflexion est le double de celle immédiatement inférieure.

Description des structures

Les formes A et B ne diffèrent entre elles que par une légère torsion et une faible rotation d'une partie de la molécule par rapport à l'autre autour de la liaison C_{1'}-N₃.

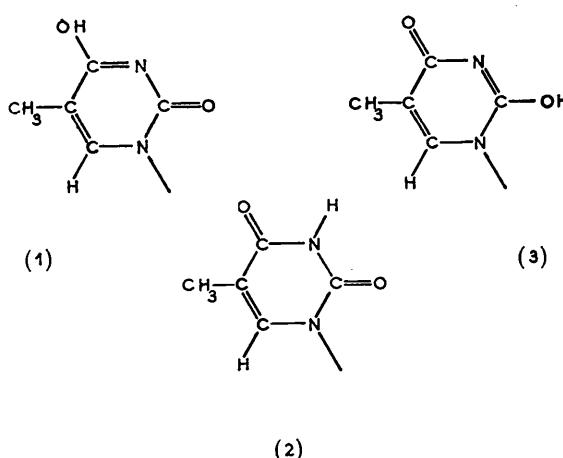
Ces deux structures confirment avec certitude que la thymidine est le 3(2'déoxy) ribofuranoside de la thymine comme il avait été indiqué par Brown & Lythgoe (1950).

Le sucre est sous forme cyclique. Les plans moyens des deux cycles sont sensiblement à angle droit, comme cela s'observe dans les composés de ce type. La configuration β de la liaison C_{1'}-N₃ se trouve confirmée. Le cycle pentagonal n'est pas plan, mais déformé de telle façon que, C_{3'} s'écartant notablement du plan moyen (Fig. 2(d)) l'oxygène O₃ s'en rapproche. La même déformation avait déjà été remarquée pour le sucrose et la cytidine (Beevers & Cochran, 1947; Furberg, 1950).

La présente étude montre que la structure de ce dérivé bromé est en accord avec la structure du reste déoxyribofuranoside supposée par Crick & Watson (1954) dans leur modèle de l'acide déoxynucléique (DNA).

Liaisons hydrogènes et formes tautomères

Le noyau hexagonal est susceptible d'exister sous trois formes tautomères:



La molécule ne porte que deux atomes d'hydrogène capables de former des liaisons hydrogène, à savoir l'hydrogène porté par O_3' et l'hydrogène labile du cycle hexagonal.

L'hydrogène porté par O_3' ne peut servir que dans les liaisons $O_3'-O_6$ (2,7 Å) pour la forme *A*, et $O_3'-O_2$ (2,8 Å) pour la forme *B*, les autres distances d'approche étant supérieures ou égales à 3,3 Å. Cette disposition sans ambiguïté d'une part, l'alignement spatial presque parfait d'autre part des atomes $C_4N_1O_3'$ rendent probable la liaison hydrogène N_1-O_3' pour les deux formes, ce qui donnerait à *A* et *B* la même formule tautomère (2).

Les deux formes A et B ne diffèreraient donc que

par une symétrie, en passant de l'une à l'autre, de liaison hydrogène par rapport à la droite $C_4N_1O_3$, formes probablement très voisines du point de vue énergétique, ce qui expliquerait assez bien la cristallisation simultanée, ou quasi simultanée, des deux formes cristallines.

Nous tenons à exprimer ici, notre reconnaissance à Sir Lawrence Bragg qui a bien voulu suivre ce travail, au Docteur Taylor pour ses bienveillants conseils au cours de cette étude et au Docteur Cochran qui nous a constamment guidé et a effectué pour nous certains calculs sur la machine EDSAC; nous remercions également le Mathematical Laboratory pour l'aide qu'il nous a apportée au cours de nos calculs.

Notre gratitude va aussi au British Council qui nous a permis de poursuivre cette étude à l'Université de Cambridge où la majeure partie de ce travail a été accomplie.

References

- BEEVERS, C. A. & COCHRAN, W. (1947). *Proc. Roy. Soc. A*, **190**, 257.

BROWN, D. M. & LYTHGOE, B. (1950). *J. Chem. Soc.* p. 1990.

COCHRAN, W. & DYER, H. B. (1952). *Acta Cryst.* **5**, 634.

CRICK, F. H. C. & WATSON, J. B. (1954). *Proc. Roy. Soc. A*, **223**, 80.

DYER, H. B. (1951). *Acta Cryst.* **4**, 42.

FURBERG, S. (1950). *Acta Cryst.* **3**, 825.

Tables Internationales pour la Détermination de la Structure des Cristaux. (1935). Berlin: Borntraeger.

Acta Cryst. (1957). 10, 133

The Structure of α (V-Al)

By P. J. BROWN

Crystallographic Laboratory, Cavendish Laboratory, Cambridge, England.

(Received 22 August 1956)

The most aluminium-rich compound in the V-Al system has been shown to be isomorphous with $Mg_3Cr_2Al_{18}$. Accurate values for the atomic parameters have been determined. The interatomic distances and some special features of the structure are discussed.

1. Introduction

The V-Al system was investigated by Carlson, Kenney & Wilhelm (1955). They found three peritectic phases α , β and γ at the aluminium-rich end of the phase diagram. The most aluminium-rich of these phases, α , was shown to be cubic with $a = 14.586 \text{ \AA}$ and to have space group $Fd\bar{3}m$; it was given the formula VAl_1 .

2 Material

The crystals used in the present work were kindly given to the Laboratory by Dr A. D. I. Nicol, for

whom they were made by Dr D. W. Wakeman. They were extracted electrolytically from a slowly-cooled ingot containing about 3% wt. vanadium. The extracted crystals were of two kinds, one of which proved to be the α phase of Carlson *et al.* These crystallized as octahedra; the other crystals, monoclinic plates having formula unit approximately VAl_7 , do not correspond to any of the phases found by Carlson *et al.*

Several analyses of the cubic α crystals were obtained. The results differed rather widely, probably because of the very small quantity of vanadium in the samples. The analysis for which the largest quantity,